2021

ESSAYS ON EVASION AND ENFORCEMENT IN VALUE ADDED TAX (VAT)

Syed Jawad Ali Shah
University of Kentucky, jawad_shah54@hotmail.com
Author ORCID Identifier:
https://orcid.org/0000-0002-3103-8584
Digital Object Identifier: https://doi.org/10.13023/etd.2021.284

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Syed Jawad Ali Shah, Student

Dr. David R. Agrawal, Major Professor

Dr. Rajeev Darolia, Director of Graduate Studies

ESSAYS ON EVASION AND ENFORCEMENT IN VALUE ADDED TAX (VAT)

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School at the University of Kentucky

By

Syed Jawad Ali Shah

Lexington, Kentucky

Co-Directors: Dr. David R. Agrawal, Associate Professor of Economics and Public Policy

And: Dr. William H. Hoyt, Professor of Public Finance

Lexington, Kentucky

2021

https://orcid.org/0000-0002-3103-8584

ABSTRACT OF DISSERTATION

ESSAYS ON EVASION AND ENFORCEMENT IN VALUE ADDED TAX (VAT)

Value added tax (VAT) based on credit invoice system is the most common consumption tax in the world. Despite its self-regulating nature, VAT faces challenges in developing countries who have limited state capacity to check evasion and enforce tax on informal sectors of the economy. The tax authorities introduce policy interventions that can target the evasive behavior of firms interacting with informal sectors. My dissertation seeks to provide insight into three such policy reforms in Pakistan's VAT regime. Therefore, this dissertation is composed of three essays.

In first essay of my dissertation, titled "Using Computerization to enforce VAT: Evidence from Pakistan", I study a policy intervention which empowered a computerized system to check invoices and reject input tax claims based on risk-based criteria. I use administrative tax data for the universe of VAT returns filed in Pakistan from tax year 2009 to 2016 to estimate the impact of this reform on the firms operating domestically. Using the exporters not subject to the reform as a control group, I find that the input tax claims fell by 2.36 million Pak Rs. per treated firm, representing a decline in input tax claims to the tune of Pak Rs. 86 billion. Firm heterogeneity analysis by business activity and firm structure shows a decline ranging from 30% to 90%. Surprisingly, the corporations and partnerships also show significant reduction in input tax claims from 50-70%. Contrary to the expectations, the huge volume of evasion shows that VAT implementation in limited tax capacity regimes may not yield the expected revenue efficiency gains.

Second essay of my dissertation titled, "Is Minimum the Maximum? Tax Burden on Informal Sector in VAT: Evidence from Pakistan", analyzes another policy reform. In developing countries, a substantial amount of revenue at import stage is now collected from VAT instead of traditional import tariffs. This modern approach assumes negligible VAT evasion at post-importation stage. I test this assumption through universe of monthly VAT returns filed in Pakistan for tax years 2009 to 2016 to estimate evasion by firms exclusively engaged in imports. I utilize kinks produced by minimum value addition thresholds to estimate evasion of VAT post-importation. I estimate an average evasion rate of nearly

78%. Using changes in thresholds over years, I provide evidence that this minimum tax collection is the best-case scenario for revenue efficiency. The firms show strong bunching at or below threshold with about 40-60% of the firms showing bunching behavior. My results support the view that, absent deviations from standard, replacing high import tariffs with VAT would decrease welfare.

Third essay of my dissertation titled, “The Deterrence Value of Tax Audits: Estimates from a Randomized Audit Program”, analyzes a randomized audit program. It is a joint project with Michael Best and Mazhar Waseem. In modern tax systems audit is the sole instrument through which the tax authority can detect noncompliance and create deterrence. We exploit a national program of randomized audits covering the entire population of VAT filers from Pakistan to study how much evasion audit uncovers and how much evasion it prevents by changing behavior. While audit uncovers a substantial amount of evasion (the evasion rate among firms in the bottom three size quartiles is more than 100%), it does not deter future cheating. Examining more than ten intensive and extensive margin outcomes, we detect no effect of audit on proximate or distant firm behavior. Our results suggest audits are sub optimally utilized in checking mechanical violations of law instead of creating deterrence against evasion.

Keywords: Value Added Tax, Tax Evasion, Enforcement, Firm Behavior, Informal Economy, Bunching

Syed Jawad Ali Shah

Date: June 21, 2021

ESSAYS ON EVASION AND ENFORCEMENT IN VALUE ADDED TAX (VAT)

By

Syed Jawad Ali Shah

Dr. David R. Agrawal

Co-Director of Dissertation

Dr. William H. Hoyt

Co-Director of Dissertation

Dr. Rajeev Darolia

Director of Graduate Studies

Date: June 21, 2021

Dedicated to my mother Fouzia Amin and my father Amin Shah, for always supporting me in my academic and professional endeavors. I also dedicate this dissertation to my wife Sidra and our beloved son Maarij.

Acknowledgments

I am grateful to my advisors David Agrawal, William Hoyt, Carlos Lamarche and Ron Zimmer for their excellent support and encouragement throughout dissertation research. Dr. David Agrawal has shown great commitment to mentoring me and I would be forever indebted to him for this. The faculty at Martin school contributed in my development in so many ways. I am thankful to Mazhar Waseem for his helpful insights and comments. I also thank Michael Best, Jennifer Bird-Pollan, John Butler, Wei Cui, Andrew Goodman-Bacon, David Hulse, Henrik Kleven, Kaisa Kotakorpi, Benjamin Lockwood, Olga Malkova, David Wildasin, and Owen Zidar for their helpful comments and advice. I am grateful to the participants of conferences, workshops and seminars held by National Tax Association, Said Business School at Oxford University, Association of Budgeting and Financial Management and University of Kentucky for providing important feedback. I also thank Amanda McCullough, Himawan Saputro, Kenneth Tester, Quinton White and participants of various workshops and seminars at University of Kentucky for their helpful advice and comments.

Chapter 3 of my dissertation is a joint collaboration with Michael Best and Mazhar Waseem. Their expertise and knowledge was very helpful in completion of this dissertation.

Table of Contents

List of Tables

List of Figures

Chapter 1

Using Computerized Information to Enforce VAT: Evidence from Pakistan

Developing countries have a very low tax to GDP ratio compared to developed countries. A key explanation is the low enforcement capacity of tax administrations in developing countries (IMF 2011). For this reason, in tax policy debates, enhancement of administration and enforcement capacity of developing countries with large informal sectors, is considered pivotal to collecting adequate taxes (Slemrod 2019; Waseem 2018; Slemrod and Gillitzer 2014). In the last several decades, over 160 countries -- including many developing countries -- have introduced value added tax (VAT). The prime motivation is the supposedly superior tax enforcement properties of VAT due to cross-checking of information across various stages of production. Each stage of production reports the value of outputs and inputs, which means that the output of an early stage of production acts as an input for the next stage of production. As these inputs and outputs are reported by unrelated firms, they create a paper trail that the tax authority can exploit for enforcement.

Contrary to the popular belief that VAT immediately provides a way of enhancing revenue efficiency, these enforcement advantages of VAT may not work in low state capacity countries. The conventional wisdom is that information flows created by arm-length transactions between unrelated parties make it easier for the government to enforce the tax. This argument ignores the ability to process the information flow and assumes that governments have the administrative capacity to utilize it. If the tax administration does not have the capacity to cross-match information and/or to enforce recovery, the enforcement-facilitating mechanism built into VAT would not be effective. Therefore, developing countries may have adopted VAT without adequate enforcement capacity to collect more revenue through VAT. In this view, and in contrast to the conventional wisdom, enforcement improvements are minimal which, in turn, implies that VAT may not be the appropriate tax for these countries to adopt (Emran and Stiglitz 2005). I study VAT in both low enforcement capacity and high enforcement capacity environments in Pakistan to answer one of the most important questions in public finance: what is the effect of third party information reporting on tax systems vis-à-vis the enforcement capacity? To preview the results: information flows alone are *not sufficient* to secure firm compliance.

Recent empirical evidence shows that firms report their sales accurately when they are more likely to be scrutinized or cross-matched (Carrillo, Pomeranz, and Singhal 2017; Pomeranz 2015). However, this enhanced probability of detection of sales may not necessarily translate to increased revenue when low enforcement capacity and legal loopholes can let firms inflate their purchases and effectively pay the same net tax.[1] Despite substantial evidence that third party reporting increases declared sales in VAT, we have little to no evidence on evasion through manipulation of purchases because, without the administrative tax records, the linkage between input claims and revenues is difficult to establish in VAT.[2] This gap in literature is very significant for VAT, which relies strongly on value addition and suffers from potential fraudulent schemes involving fake invoices which are not found in other taxes. VAT regimes can have bogus traders who only register to serve as "invoice-mill" and generate fake invoices (Waseem 2020). Tax evading firms use these fake invoices as input tax credit to lower their tax liability. Even in the developed countries such as those in EU, Missing Trader Intra Community (MTIC) or "Carousel" fraud, is rampant. EU almost gave up the destination based taxation principle as it could not cope with large volume of revenue leakage through carousel fraud (Crawford, Keen, and Smith 2010; Keen and Smith 2006). These sophisticated and unique VAT only frauds can only be tracked through a real time system which can audit and recover tax in a fast and efficient manner.[3]

In this paper, I exploit quasi experimental setting created by a Pakistani reform which authorized a software based risk analysis system named CREST[4] to deny suspicious input tax claims in real time. This reform discretely raised the ability of the government to utilize information flows. CREST has access to data other than VAT

[1]Carrillo, Pomeranz, and Singhal (2017) provide an interesting case of this phenomenon using the firms in Ecuador who reported increased sales when faced with the prospect of detection but they also increased costs 96 cents to a dollar to effectively wipe out any revenue impact.

[2]Waseem (2018)provides evidence on increased sales in response to enforcement but also shows that there is little to no effect on informal sectors who remain insular to the VAT. Fan et al. (2018) show that invoice summary cross-matching increases revenue.

[3]Annually, more than hundred billion Euros worth of taxes are lost in EU. Therefore, it's not hard to imagine the scope of difficulties faced by the revenue administrations in the countries with large informal economies in curbing fraudulent practices. Interestingly, in the countries with large informal economies these frauds become more easier to execute because of lax enforcement.

[4]CREST stands for Computerized Risk-Based Evaluation of sales tax. In Pakistan, VAT is legally called sales tax because this law was introduced as an amendment to existing statute of sales tax and not as a new law to comply with the constitutional requirements.

returns, can go few steps back in chain[5] and uses in-built risk parameters to establish the authenticity of each and every invoice. It reduces the role of the "*taxman*" (auditors) and replaces these traditional enforcement mechanisms-- plagued with inefficiency, corruption (Davoodi and Tanzi 2000; Khan, Khwaja, and Olken 2016) and delays-- with an efficient, transparent and real time enforcement system. In particular, starting July 2013, CREST software rejects input credits automatically by performing an invoice wise scrutiny. The reform eliminated the need to adopt a long and tedious process starting from audit selection, completion, framing a case based on audit, ensuring "over the years" that the case reaches its logical end and, critically, the tax is recovered from the defaulting unit. It took away the opportunity from the tax evaders to use various loopholes in this process.

I develop a general conceptual framework for input tax evasion in VAT in the spirit of model of tax evasion introduced by Allingham and Sandmo (1972). In this standard model, the probability of detection is represented in a reduced form way as a single parameter. However, the probability of detection depends on the product of the probability of audit and the probability of *recovery* conditional on being audited. The CREST reform raised the expected cost of evasion by substantially increasing the probability of *recovery*. I, therefore, use two different probabilities in my theoretical model to capture this effect of the reform. From theoretical model, I conclude that unless governments invest in improving the probability of recovery, increasing the number of audits would not deter evasion.

Turning to the empirical analysis, critically, the reform affected non-exporting firms only as exporting firms were already subject to virtually identical scrutiny prior to the reform, thereby facilitating a generalized difference-in-differences identification strategy. I use administrative data[6] for the universe of monthly VAT returns (9.9 million in total) filed in financial years 2009 to 2016[7] to study the input tax based evasion. Using exporting firms as the comparison group and domestically operating firms as the treatment group, I find that the input tax claims fell by 2.36 million Pakistani Rupees per treated firms, which represents a decline of 50% on average for the treated firms. Using a generalized difference-in-differences design, I provide

[5] It means the software can cross check invoices and then checks suppliers of the supplier and so on which enables it to crunch data and raise a red flag immediately. It can also cross check import and export records, Income tax returns and excise tax etc.

[6] See also Waseem (2020), Slemrod, Ur Rehman, and Waseem (2020), Waseem (2019), Waseem (2018), Best et al. (2015) and H. Kleven and Waseem (2013) for recent empirical evidence using administrative data from Pakistan.

[7] In Pakistan financial year starts on July 1st and ends on June 30th, accordingly the tax year 2009 means the financial year starting on July 1, 2018 and ending on June 30, 2009.

evidence to support the parallel trends assumption by finding that exporting and domestic firms show nearly perfect parallel trends in the pre-reform period . These large post-reform effects are not limited to or driven by individually operated firms. I find that corporations and partnership firms also show similar decline in input tax claims ranging from 30-50%. My results provide evidence of significant VAT evasion under the regime with traditional ("by hand") audits. It implies that the self-enforcing advantages of VAT do not hold in absence of effective enforcement which can utilize third party reporting. Although, introducing digitization and computerization – and thus increasing enforcement capacity – to cross-check available third-party information and build a risk profile for every credit invoice substantially reduces tax evasion but it also poses a larger question about efficacy of VAT in developing countries. Absent a comprehensive real time enforcement mechanism which can utilize these information flows and automatically deny credit up front, the volume of evasion would remain very high. Returning to the big picture public finance question, my evidence suggests that advocating VAT adoption in low-tax capacity environments may have been premature. In particular, using all the available information trails to enforce VAT is not possible unless the government has high tax capacity.

I divide my analysis into three parts. First, I determine the extent of evasion through manipulation of input tax credits by estimating the drop in total input tax credits which is approximately 50%. Second, I estimate the impact by business type (company, sole proprietorship or partnership) which shows that the companies, who would otherwise be expected to refrain from outright fraud, show behavior similar to partnership and sole proprietorship firms. Third, I measure the effect across business categories (manufacturers vs. non-manufacturers) which ranges from 30% for the manufacturers to 70% for the non-manufacturers.

This paper adds to three different strands of literature. First, it adds to the literature on enforcement capacity of the developing countries which have large informal sectors and its implications for welfare effects of VAT.[8] Main focus of this literature is sales and output tax with very less emphasis on purchases (Naritomi 2019; Pomeranz 2015; Slemrod and Gillitzer 2014; Crawford, Keen, and Smith 2010; Paula and Scheinkman 2010; Piggott and Whalley 2001). Best et al. (2015) use Pakistani

[8]Emran and Stiglitz (2005) had argued that a revenue neutral VAT would actually decrease welfare in developing countries with large informal sectors. This view relies on an expectation that evasion in developing countries would remain high. Keen (2008) countered this argument partially by arguing that an increased collection of VAT on import stage ensures that the informal sector is sharing the tax burden. My paper provides further evidence for the former argument by showing that the tendency of formal sector to profit from fake invoices against purchases made in informal sector can be high.

income tax data to show that corporate revenues increase from a turnover tax compared to standard profit and loss based income tax because, in the latter case, firms can manipulate their purchase figures to lower their liability substantially. Similarly, Carrillo, Pomeranz, and Singhal (2017) use Ecuadorian data to show that the firms inflate their purchases to effectively offset any gains received from truthful sales reporting. I add to this purchase inflation strand of the literature by providing evidence of purchase manipulation and the efficacy of third party information in enforcement using administrative data.

Second, this paper contributes to the literature on the effectiveness of invoice summaries using administrative data and examines effect of second stage verification of invoices.[9] This literature only focuses on first stage verification in which effect of invoices verification through cross checking only is estimated. Waseem (2020) studies the self-regulating nature of VAT using Pakistani tax returns data but the major period covered in the paper is before the introduction of invoice summaries in Pakistan and does not deal with the impact of invoice summaries. Fan et al. (2018) study the impact of invoice summaries in China but they do not use actual returns data and Chinese VAT is not a standard VAT implemented in Pakistan and elsewhere.[10] However, my paper not only provides the evidence of effectiveness of invoice summaries using actual return data but also examines the impact of building a risk analysis system which integrates invoice summaries with other information available to the tax authorities to preempt evasion and fraud.

Third, the paper adds to the literature concerning destination versus origin based commodity taxation and their impact on evasion (D. Agrawal and Mardan 2019).[11] The standard VAT which relies on destination based taxation, provides opportunities for evasion and frauds where presence of multiple jurisdictions can be used to circumvent the ability of tax administration to detect evasion and recover the tax. I also add to a wide literature on missing trader fraud which, however, lacks in empirical

[9] First stage verification means that invoices are only cross checked to ensure that any credit claim is based on an actual invoice. This verification is analogous to verification of income tax withholding certificate issued by employers and contractors in which only the amount of withholding is cross checked. Second stage verification of VAT credit invoices deals with ascertaining that the transaction reported to tax authorities by buyer and seller actually took place or it was only part of a potential fraudulent activity. I explain this mechanism in detail in Appendix A.1.1.

[10] Fan et al. (2018) study the impact of computerized invoices on Chinese manufacturing firms for the period 1998-2007. They find that computerization explains 14.38% of cumulative VAT revenues and they also find a 4.7-14% increase in the average effective tax rate for seven subsequent years of the reform.

[11] D. Agrawal and Mardan (2019) provide the simplest understanding of the issues involved in destination versus origin based commodity taxation using the tax avoidance and evasion behavior in sale and use taxes levied by sub-national jurisdictions in USA.

studies. I provide empirical evidence on the prevalence, dynamics and working of missing trader fraud in VAT on an otherwise vast literature on carousel fraud. The problem of VAT evasion through fake invoices is so widespread and voluminous that it has often prompted EU to move away from a standard destination based VAT (see Fath, Goulder, and Williams 2015; Keen and Smith 2006; Bickley 2003).[12] The reform worked and can be modified to any other country or setting facing rampant missing trader fraud. Many variants of VAT such as a compensating VAT (CVAT) and Viable Integrated VAT (VIVAT) to deal with intra community and sub-national implementations, have been proposed (Bird and Gendron 2000). The paper shows that a standard destination based VAT can utilize computerization for real time verification to improve compliance and prevent fraud.

This paper calls into question proliferation of VAT in developing countries based on the premise that self-enforcing nature of VAT would automatically translate to revenue gains in these limited enforcement capacity regimes. In the absence of an end-to-end automated system, risk based analysis and enforcement in real time, the magnitude of evasion would remain very large and the expected gains in revenue efficiency would not materialize. Therefore, there is a reason to believe that VAT implementation in developing countries happened at least a couple of decades earlier and without building the necessary institutional capacity.

1.1 Institutional Setting

To understand the context of Pakistani reform, I first elaborate fake invoice phenomenon in VAT. Typically, fake firms register with VAT administration and without actually carrying out any business activity, they issue invoices which can be later claimed by the operating units. These fake firms then disappear without remitting the tax due to the government. Therefore, this type of fraud is often called "missing" trader fraud. The operation and extent of the MTIC fraud varies from one VAT regime to another but the central idea is same. A group of traders purchases and sells goods between themselves in a manner that one or several of them vanish without remitting the tax collected, thereby forcing tax authority to allow credit for the

[12] Keen and Smith (2006) elaborate this phenomenon. Annually, more than hundred billion Euros worth of taxes are lost in EU. Therefore, it's not hard to imagine the scope of difficulties faced by the revenue administrations in the countries with large informal economies in curbing fraudulent practices. Interestingly, in the countries with large informal economies these frauds become easier to execute because of lax enforcement (see also Bickley 2003; Bird and Gendron 2007).

amount which was never deposited in the first place.[13]

In European countries, the carousel fraud relies on trade within EU because the tax administrations do not collect tax at import stage on imports originating from member countries (see Figure 1.1).[14] But in most developing countries, who charge tax on every import without any exception to a particular origin, domestic variants of carousel fraud may exist (Keen and Smith 2006). Consequently, the tax authorities come up with a variety of enforcement and legal measures to curb this phenomenon (Crawford, Keen, and Smith 2010). I refer to this type of fraud as "Domestic Missing Trader" or DMT fraud. Pakistan has a large informal sector providing significant incentive to claim input tax credit against purchases, which are actually made in informal sector. Fake invoices have a readily available market which wants to show purchases from informal sector as purchases from VAT registered firms to claim excess credit. I explain this phenomenon in the context of Pakistani VAT regime (see Appendix A.1.1 for details).

1.1.1 Legal Framework and the CREST Reform

VAT is the principal source of revenue for the Federal Government in Pakistan and FBR administers VAT. The governing legislation is an Act of Parliament hereinafter referred as "the Act". The Act allows executive branch to make rules which provide the administrative framework to implement VAT. These rules lay out the administrative procedures such as the registration rules which govern registration and deregistration of firms. The "Refund" rules are also part of this statute and outline the mechanism for filing, processing and sanctioning of refund claims against zero rated (mainly export) supplies. The bulk of refund claims, more than 97% in value, relate to exports. Under these "Refund" rules, firms file their monthly claim electronically and provide supporting documentation to the concerned refund processing division. CREST cross matches the information provided with refund claim including purchase and sale invoices with the data available in the system. It generates risk based assessment on each purchase invoice pointing out the type and nature of discrepancy. It explicitly states whether an invoice is "valid" or "invalid" along with the reason (see Appendix A.1.1 for more details on CREST).

[13]The VAT invoice is similar to an income tax deduction certificate for a buyer. The invoice shows that tax has been collected by seller and buyer is now entitled to deduct this amount from any sales made during the same tax period. If the tax deducted at purchases is more than the tax required to be deposited on sales then firms can either carry it forward to the next period or seek refund.

[14]Keen and Smith (2006) elaborate the operation of carousel fraud in EU. Figure 1.1 is adapted from the discussion in their article.

Ironically, no provision was available in any law to apply information obtained from CREST to check the firms who are not claiming refund prior to financial year 2014. This implied that as long as the firms did not claim refund, tax authorities had little room to check proactively whether the tax against which input is being claimed had actually been deposited in treasury by the supply chain of domestic supplier. Furthermore, there was no legal cover because the rules applied only to refund claimants and any proceeding against domestic units based solely on CREST would be legally void. The usual mechanism of selection of audit and the pace on which the audit proceeds meant that the network of fraudsters could go unchecked for years causing staggering losses to the exchequer. The absence of a legal cover and the lack of administrative impetus to check this phenomenon in real time meant that the refund claimants and non-claimants were essentially operating under two different audit and enforcement regimes. The revenue cost of a fake invoice is same for the tax authority in either case but the firms who did not file refund claims could only be caught through an audit. FBR had the information since 2008 through CREST but it had no meaningful way of using this third party information to prevent evasion in the domestic cases as it did in the cases related to exports. Selection for audit is a very low probability outcome compared to the compulsory transaction level scrutiny required for a refund claimant.

From 1st July 2013, through a change in the Act, the legislature made objection raised by CREST a valid criterion to reject input claim.[15] This implied that tax administration could reject the input tax credit of non-refund claimants proactively and initiate proceedings using the information obtained from CREST. Instead of the low probability selection for audit, the domestic suppliers relying on fake invoices now faced a real time challenge. The long and tedious process of audit selection and recovery which took years previously could now be done instantaneously. Suspicious supply chain became a valid ground of rejection through CREST and domestic firms as well as exporting firms faced identical scrutiny post reform should they choose to evade. FBR forcefully implemented the reform by introducing instructions to administrative units to check input invoices on monthly basis and point out the discrepancies. Tax authorities could now generate an automatic notice of rejection of input tax credit and contest the case legally because the Act now provided cover to this real time rejection. FBR could check the networks of fake suppliers and input tax credit could

[15] The law does not require a further proof from the tax authority. If CREST terms an invoice "invalid" then the onus shifts to the firm to satisfy the objections raised. Most importantly, CREST can raise objection based on suspicious supply chain against an individual invoice or transaction and reject input tax credit involved therein

be denied by the software without the involvement of the taxman.

In short, the reform discretely increased the enforcement capacity from low to high. Prior to the reform, FBR could detect the fraud but statutory limitations on audits and the traditional "by hand" audit approach which is plagued with inefficiency, corruption, lack of resources and legal loopholes, failed to translate in meaningful recovery conditional on detection. The reform substantially increased the probability of recovery conditional on detection. It transformed a low enforcement capacity VAT regime into high enforcement capacity regime for the domestically operating firms. The reform also affected the cost of obtaining fake input tax invoices, albeit indirectly, by shortening the period within which detection takes place.[16]

1.2 Theoretical Model

I develop a model of input tax evasion based on model of Allingham and Sandmo (1972), hereinafter referred simply as A-S model. Although this model is based on income tax evasion, but the intuition employed in A-S model is applicable to the present case of input tax evasion in VAT. The tax evasion literature on developing countries has widely used A-S model (for example-Carrillo, Pomeranz, and Singhal 2017), but only for output tax evasion. I use the basic intuition in this model and modify it to input tax evasion. A-S model uses only one parameter for the probability of detection, which includes the probability of recovery conditional on detection. In VAT, however, probability of detection and the probability of recovery conditional on detection can vary differently. The reform only changed the probability of recovery conditional on audit. Therefore, I use separate parameters to capture the effect of detection and recovery, conditional on audit. This change would not only decrease input tax claimed by manufacturing units who were using fake input tax credit but also substantially reduce both input and output tax of the fake units which were previously churning out fake invoices with impunity. Consequently, my model predicts large effect of the reform on input tax evasion. The net tax gain to the government, though substantial, would be less than the total fall in input tax credit claimed because a good chunk of this observed drop would be driven by fake units.[17]

---

[16] This requires an understanding of the black market where these invoices are sold on a cost proportional to the fake credit involved which is discussed in detail later. However, in short, the suppliers of fake invoices had a greater risk that their invoices would be caught before an input tax is claimed by the beneficiary which meant that the suppliers of these invoices would receive zero payment (and even retribution!) plus loose whatever they have already invested.

[17] Keen and Slemrod (2017) suggest an alternate way to model the response by using the elasticity of tax revenue with respect to an intervention. It would be empirically difficult to determine this

Firm's choice to evade relies on a simple decision. If the expected benefit of claiming excess input tax exceeds the expected cost then the firm has an economic motivation to evade. Consider a firm which has taxable output y, and a taxable input x. For simplicity, I assume that both input and output are taxed at uniform rate τ. Input tax can be divided into two parts based on whether a legitimate VAT invoice is available for that or not. Therefore, x is composed of two components x_1 and x_2 which represent the real taxable input and the fake taxable input respectively. Then firm's VAT liability for a certain tax period is given by: $Z = (y - x_1 - x_2)\tau$. I denote the firm's actual tax liability, $(y - x_1)\tau$ by $Y$and fake input $x_2\tau$ by F. In case there was no restriction or cost to evasion then each firm shall report $Y = F$ so that its VAT liability $Z = 0$. If the firm is generating an income,W, then any fake input tax claim adds to its income.

A-S model of evasion is based on probability of detection, p, through an investigation. In income tax, if the tax authority detects undeclared income then you have to pay tax on undeclared income. In limited state capacity regimes, relation between detection and recovery is not straight forward. The firm which relies on fake invoices often gets away with fraud because the "shady" link between buyer and seller is difficult to prove in courts. Tax authorities have to credit the input claimed in fraudulent manner because they are unable to trace or prove the case against fake suppliers. Therefore, detection does not automatically translate to recovery. I, therefore, use two different probabilities to capture this effect of the reform. Firms are aware of the loopholes and would take into account probability of detection, p_1, as well as recovery, p_2.

The expected cost of evasion is composed of three components: a) cost of obtaining fake input tax invoices (b) the recovery in case of detection which includes penalty (c) legal fees associated with audit and litigation incurred by the firm whether the revenue authorities make or fail to make any recovery. I use separate parameters for the cost components associated with traditional legal fees and the cost of obtaining a fake invoice. The penalty,π, is proportional to the tax evaded.[18] Similarly the cost of obtaining fake invoices, θ, and the legal expenses incurred, l, are also assumed proportional to the tax evaded.

elasticity in this case because of this missing trader response.

[18]The modification of A-S model given byYithzaki (1974) uses a penalty rate proportional to the tax evaded, and in most countries including Pakistan, this is a standard practice.

The firm will choose F to maximize the expected utility given by:

$$E(U) = \underbrace{(1-p_1)U(W+F-\theta F)}_{\text{No detection}} + \underbrace{(p_1 - p_1p_2)U(W+F-\theta F - lF - l\pi F)}_{\text{Detected but not recovered}} \\ + \underbrace{p_1p_2U(W+F-\theta F - lF - l\pi F - \pi F)}_{\text{Full recovery including penalty}} \quad (1.1)$$

where $0 \leq \theta, l, p_1, p_2 \leq 1$. The limits on probabilities are obvious. Value of θ greater than 1, shall imply that the cost of obtaining input invoices, before a return is filed, is more than the tax involved in those invoices. Similarly, legal fees cannot be more than the actual tax plus penalty demanded because the firm would then simply pay the amount detected. For notational convenience, I denote functional terms other than W in (1) by $G^a, G^b$and G^c where

$$G^a = F - \theta F, G^b = F - \theta F - lF - l\pi F, G^c = F - \theta F - lF - l\pi F - \pi F \quad (1.2)$$

so that

$$E(U) = (1-p_1)U(W+G^a) + (p_1 - p_1p_2)U(W+G^b) + p_1p_2U(W+G^c)$$

and the first and second order conditions are then

$$(1-p_1)(1-\theta)U'(W+G^a) + (p_1-p_1p_2)(1-\theta-l-l\pi)U'(W+G^b) \\ + p_1p_2(1-\theta-l-l\pi-\pi)U'(W+G^c) = 0 \quad (1.3)$$

$$(1-p_1)(1-\theta)^2U''(W+G^a) + (p_1-p_1p_2)(1-\theta-l-l\pi)^2U''(W+G^b) \\ + p_1p_2(1-\theta-l-l\pi-\pi)^2U''(W+G^c) = 0 \quad (1.4)$$

The second order conditions are satisfied because the utility function is concave. The interior maxima shall exist between $F = 0$ and $F = Y$ but only subject to $\theta = l = 0$.[19]Since As expected marginal utility is increasing in F, evaluating eq.1.3 at

[19]In VAT, a firm can "theoretically" claim as much fake input in a tax period as it wants because the excess will be refunded. But the finite limit on a long term horizon is imposed by the positive value addition factor, therefore Y=F is a justified upper limit on F with $\theta = l = 0$.

these two points result in following two relationships.

$$\frac{\partial E(U)}{\partial F}\mid_{F=0}= (1-p_1)(1-\theta)U'(W) + (p_1 - p_1p_2)(1-\theta-l-l\pi)U'(W) \\ + p_1p_2(1-\theta-l-l\pi-\pi)U'(W) < 0 \quad (1.5)$$

$$\frac{\partial E(U)}{\partial F}\mid_{F=Y}= (1-p_1)(1-\theta)U'(W+(1-\theta)Y)+(p_1-p_1p_2)(1-\theta-l-l\pi)U'(W+(1-\theta-l-l\pi)Y) \\ + p_1p_2(1-\theta-l-l\pi-\pi)U'(W+(1-\theta-l-l\pi-\pi)Y) > 0 \quad (1.6)$$

The conditions from 1.5 and 1.6 can be rewritten as

$$\pi p_1 p_2 < \frac{1-\theta}{1+l} - \pi p_1(1 - \frac{1-\pi}{1+l}) \tag{1.7}$$

$$\implies \frac{1-\theta}{\pi} > lp_1 + \pi p_1 p_2 + l\pi p_1 p_2 + l\pi p_1 \tag{1.8}$$

$$\pi p_1 p_2 < p_1 p_2 + (1-p_1p_2)\frac{U'(W+Y)}{U'(W+(1-\pi)Y)} \tag{1.9}$$

The term on the right side of 1.9 and 1.8 are positive and less than one. Therefore, 1.7 and 1.9 together give positive parameter values which are sufficient for an interior solution. I use the equations derived above to model the response of a firm to DMT. The revenue authority wants to increase the cost component or θ. If they increase the cost of registering a new bogus firm, it can impact the fraud but at the same time it is very difficult to deny registration to a business on the basis of a presumption. This would create more difficulties for genuine businesses and hence should be ruled out as a possibility. However, the converse may be true here. The tax authority would minimize the cost of registration. The "invoice mills" would charge the beneficiary units, a fixed percentage of the tax involved in fake invoices. The legal fees are determined by the market and the government has no control over it, but for the firm they also come at a cost which is proportional to the tax and penalty demanded in audit observations. Theoretically, tax rate τ, penalty rate π, increasing p_1 through more audits, and ensuring recovery after detection thereby increasing p_2, are the only options available to revenue authorities. As tax rate τ decreases, benefit of evasion also decreases and cost component dominates but small tax rate cannot generate adequate revenue. This rules out major decrease in tax rate. In Pakistan, penalty for tax fraud is 100% of the tax evaded which means $\pi = 1$. This scenario implies that the product $p_1 p_2$ should be sufficiently small for evasion to occur, which means that if either $p_1$or p_2 is small the missing trader fraud becomes economically feasible. Because it is easier for the government to detect fraud after a certain interval of time, the p_1 factor remains relatively high. In fact it is the inability of an enforcement regime to recover tax post detection (low p_2), which provides an environment conducive to this type of fraud. Missing trader fraud in Europe exploits the lack of sufficient inter country coordination or low p_2. Similarly, DMT in Pakistani case relies on legal loopholes and complexity of territorial jurisdictions which make the post detection recovery, a very low probability event. CREST reform has only raised p_2, or the probability of recovery conditional on detection. The department had all the information to conduct audit and frame the case before the reform. The cost of generating fake invoices θ, the legal fees l and the penalty on fraud π also stayed the same.[20] In Pakistan, the penalty on fraud is 100%, so$\pi = 1$ means 1.7 reduces to:

$$p_1 p_2 < \frac{1-\theta}{1+l} - p_1 \tag{1.10}$$

As the probability is always a non-negative number, therefore, $p_1 p_2 > 0$ which would imply $p_1 < \frac{1-\theta}{1+l}$ is the minimum condition for input tax evasion to occur. In other words, this condition can be expressed as:

$$\text{Probability of detection, } p_1 < \frac{\text{Net benefit per unit fake input tax}}{\text{Gross benefit plus legal cost per unit fake input tax}}$$

and by rewriting 1.10,

$$\text{Probability of recovery, } p_2 < \frac{\text{Net benefit per unit fake input tax}}{\text{(Gross benefit plus legal cost per unit fake input tax)*}p_1} - 1$$

The behavior of the businesses would differ based on their particular characteristics because of the factors discussed above.

1.3 Data

A major contribution of this paper is to analyze fake invoices and VAT fraud using rich administrative return data. To the best of my knowledge, no paper has studied

[20]The cost of generating invoice remained the same. The real time cross verification and audit makes it difficult for the chain of suppliers to operate because CREST detects suspicious chain before the beneficiary unit claims fake invoice. But, CREST itself does not impact cost of registering a new unit and supplying invoices.

missing trader or other VAT frauds using administrative data. I use the administrative return data for full universe of VAT returns filed for financial years 2009 to 2016. Because returns are filed on monthly basis, data provides a rich number of pre and post periods and a total set of 9.69 million observations. The data covers each field in the return which gives more than hundred variables (see Appendix A.1.2).[21] Return columns capture all possible purchase and sales transactions and clearly separate zero rated, reduced rate, special and exempt transactions etc. I can use this information to observe a firm's response by disentangling possible confounding transactions which is not possible in absence of returns data.

The variable of interest here is the domestic input tax credit claimed by firms. Domestic input tax credit arises from domestic taxable purchases only and does not include input tax credit from direct imports. Total input tax is a sum of domestic and imported input tax credit. In case, monthly input tax credit exceeds output tax then firm can either claim refund or carry forward this input tax to the next period.[22] Table A.1 presents descriptive statistics for VAT returns on a financial yearly basis. Number of returns filed increases each year which represents the entry of firms in VAT regime. Exit of firms cannot be inferred from the data as firms that apply for deregistration have to file nil returns for six months and should complete a deregistration audit with the department which renders official deregistration a costly business for both the firm and the tax administration without much benefits. Therefore, firms who are no longer in business either stop filing returns or keep filing nil returns long after they have gone out of business. The non-active or dormant units should not be part of my analysis but the return data includes such units because many businesses obtain registration and then fail to translate into an actual operative firm. The long time and costs associated with deregistration can force these firms to file returns without actually showing any activity. Furthermore, I focus in this paper on input tax credit, therefore, I drop the firms which claim total domestic input tax credit of less than PKR 10,000 (which equals \$100) over the course of five years before the reform. This criterion automatically drops inactive firms, commercial importers and any other firms which never claimed any substantial domestic input tax credit from both comparison and treatment groups. Table 1.1 provides descriptive statistics for the three categories along with their mean domestic monthly input tax credit and

[21] For the importance of administrative data in measuring compliance and enforcement, see Slemrod (2016).

[22] Although there's a bar on adjusting more than 90% of input tax against the output tax but this restriction does not apply to the wholesalers, wholesalers-cum-retailers and distributors vide S.R.O. 647(I)/2007 dated 27th June, 2007.

the firms excluded from the analysis, clearly have a negligible domestic input tax credit.

Some firms file returns on quarterly basis, therefore, I use quarter as my time period of analysis because the data does not differentiate between quarterly and monthly returns and tags it to the month for which return is filed. The quarterly returns are filed for the quarters ending in March, June, September and December. Therefore, monthly time periods would inaccurately inflate the figures for months of March, June, September and December. After converting to quarters and dropping observations as explained above, I perform main analysis on 2.35 million observations in terms of quarters, but the results are robust to alternate specifications (see Table A.2, and Table A.3).

1.4 Empirical Strategy

I use difference-in-differences (DID) design to study impact of the reform. It requires two key assumptions. First, a suitable comparison group is available to study the change. Second, the reform is exogenous such that the only change affecting treatment group is the policy intervention itself and neither treatment group nor comparison group changes its behavior in anticipation of the reform. The reform is a law change introduced by the legislature in budget[23] and plausibly exogenous keeping in view how the budget process works in Pakistan. FBR prepares revenue budget under secrecy and finance bill is only unveiled when Finance Minister introduces it in legislature, in the first or second week of June. Legislature passes finance bill before 30th June because it is applicable from the first day of July. In present case, this process effectively rules out any behavioral change after the reform is announced and before it is implemented. Moreover, the reform does not restrict CREST's scrutiny to a particular cut-off date. Limitation clause restricts show cause notice to five years, therefore, FBR can raise objections to anything which is not more than five years old. Firms would not get any benefit from changing behavior in anticipation of the change. FBR can raise a demand on invoices claimed earlier because the statutory period of limitation is five years. Therefore, there is no benefit for any firm to claim more input tax credit in anticipation of the reform. CREST was operational for five years and FBR was already using invoice data to raise audit observations against the fraudulent units, to blacklist and suspend registrations etc. but the recovery of evaded tax was a low probability outcome. Additionally, CREST was applicable to refund

[23]The budget for financial year 2014 was presented on 12th June, 2013 and the reform was applicable from 1st day of July 2013.

claimants for at least five years prior to the reform, which makes refund claimants an appropriate comparison group for DID design. An ideal DID design also assumes that the reform does not affect comparison group and only treatment group experiences effect of the reform. As CREST was already applicable to refund claimants, there is little chance that they shall be affected by the reform.

It is pertinent to note here that a firm is in comparison group if they claimed a total refund of input tax credit for the period July 2008 to June 2013 in excess of one million PKR[24] . I select this threshold because I do not expect a firm to claim refund through CREST if the total amount claimed over a period of five years is less than 10,000 US$.[25] This threshold also helps exclude the refunds which do not arise on account of exports or any zero rated supplies. Although some of these non-export refund claims, such as the refund arising out of input tax not adjusted in the relevant tax period, are processed through CREST but they do not require normal cross matching done for zero rated supplies.

I address three possible concerns over suitability of refund claimants as comparison group. First, Can the exporters be a good comparison group for domestic firms? There is an important difference between a refund claimant on account of exports and a totally exporting enterprise. The refund claimants can be firms who carry out most of their sales to domestic firms but still claim refund on the portion related to their exports. Moreover, even if the exports increase or decrease disproportionately, it would result in a corresponding increase or decrease of sales to these exporting firms by domestically operating suppliers, thereby inducing a similar economic trend in the treatment group. Second, Can the exporting firms which are larger in size with higher mean input tax credits have different attributes which materially confound identification? The exporting firms are definitely bigger firms on average but this in fact makes them better comparison group for manufacturers which also have larger size. Moreover, as discussed above, there is no reason for a larger firm not to take advantage of loopholes in enforcement differently in a VAT regime (See Waseem 2018; Pomeranz 2015). Third, Can an already treated group be a good comparison? Kotchen and Grant (2011) use a natural experiment in Indiana to study the effect of Daylight Saving Time (DST) on electricity consumption by difference in differences method. They use DID approach when some counties were always treated (had DST)

[24] The exchange rate for Pakistani Rupees to US Dollars was approximately 100:1 on 1st July 2013, therefore for ease of reference PKR 1 million translates to 10,000 US $.

[25] Although some firms may not be present for all eight years and more importantly in the years prior to reform but still it is realistic to assume that they would not venture to go through CREST voluntarily if the benefit is economically low to negligible compared to a high compliance cost.

to the counties which were compulsorily switched to DST by the state in 2006. They argue that once a group that was treated way back in time period such that it can be assumed to be always treated then DID can measure the causal effect of policy change by making it a comparison group. Same analogy fits here because CREST was applicable to refund claimants at least 20 quarters before this reform. I plot all the graphs with raw data and lead of 19 quarters to show that the trends are parallel (see figure 1.2 and figures in Appendix A.1, A.2) . Thus, in absence of the reform, trends should stay parallel. Figure 1.2 also shows that after a dip attributed to the reform, trends again become parallel albeit with a higher differential. This after trend substantiates the common shock assumptions for treatment and comparison groups.

An additional concern could be that the firms are switching from domestic purchases to imports during the aforesaid period. To rule out this possibility, I plot ratios of domestic input tax credit and imported input tax credit to the total input credit claimed and also the ratio between two types of credits in figure 1.3 to show that the trend is solely driven by the reduction in input tax without a change in imported input tax credit claims. I also control for imported input tax in all the regressions and they all have statistically non significant coefficients for imported input tax with point estimates which are also close to zero.

My analysis follows a simple difference-in-difference design at firm level with time and firm fixed effects. The equation of interest can be written as:

$$Y_{it} = \alpha_0 + \sum_{j \neq k} \delta_j (treated * I(t = j)) + \phi_i + \psi_t + X'\gamma + \epsilon_{it} \tag{1.11}$$

The dependent variable Y_{it} denotes domestic input tax credit for a firm in a given quarter; δ_js are the coefficients on the interaction dummy for all the quarters excluding the first quarter before the change and they track evolution of trend over time; ϕ_i and ψ_t are firm and quarter fixed effects respectively and γs are the coefficients on control variables[26]. The above equation is a generalized form of difference-in-differences and δ_j for all$j < k$(last quarter before the reform) capture the placebo effect for all pre time periods included in the analysis. Similarly, all $j > k$ would capture the evolution of trend over time in post reform period.

[26]I only use imported input tax credit as control variable for the regressions plotted in figures and tables.

1.5 Results

Prior to the reform, FBR had all the information but could only proceed against suspected firms if they were randomly selected for audit, a very low probability outcome as no more than 1-5% firms are audited by the department in a given fiscal year. Moreover, FBR publishes this audit schedule on its website which can easily forewarn these suppliers and they can just go missing or become non-compliant thereby denying the department any meaningful audit. The other option is to institute a criminal proceeding through an investigation but the criteria for obtaining this approval and finalizing the proceedings are also strict and time consuming. In practical terms, even with a high p_1, p_2 remains very low which keeps the term on left side of 1.10 very small. The reform increases p_2 substantially by authorizing a software to accept or reject the tax credit.

I restrict my main regression based analysis to four pre and post quarters to guard against the firm behavioral changes over time as I take firm fixed effects with standard errors clustered at firm level. To address the concerns on parallel trend assumption, I plot the interaction dummies with their 95% confidence intervals in figures 1.4 to 1.9. I omit the reform quarter dummy to avoid perfect collinearity. The lead coefficients are statistically and economically zero but there is a significant change post reform. The results are similar for both balanced and unbalanced panel. Tables 1.2 to 1.7 show the regression results for overall regression estimates and firm heterogeneity by manufacturing and business type. All regressions discussed below control for the imported input tax. The increased probability of recovery results in a sudden and lasting drop in input tax claims across the board.

1.5.1 Overall Impact

The reform decreased input tax claims by PKR 2.22 million on average for the unbalanced and PKR 2.36 million for the balanced panel. This amounts to a decline of 50% compared to pre reform levels for the balanced panel. Figure 1.4 plots the coefficients of interaction dummies (post June-2013*domestic suppliers) for the specification at equation (8) for all the firms. All interaction coefficients are close to zero and statistically non-significant pre reform but are significant post reform. It shows that the comparison and treatment groups had parallel trends prior to the reform. After the reform, input tax credit claims fall significantly for both balanced and unbalanced panels. Coefficient for the balanced panel is slightly higher (approximately 5%) than the unbalanced panel. The balanced panel comprises of firms who filed returns in all the nine quarters starting from April-June 2012 to April-June 2014. Therefore, balanced panel has only those firms which operated both before and after the reform. On the other hand, unbalanced panel has all the firms before and after the reform whether they filed returns in all quarters or not. As discussed above, entry is always more than exit which implies that more firms enter in the unbalanced panel after the reform when compared to number of firms who exit before the reform. Hence, coefficients on unbalanced panels are slightly lower than the balanced panel coefficients but they do not differ much in magnitude.

I observe two effects here, a "sudden" drop and a large change in magnitude of domestic input tax claims. These results show that the reform was effective in blocking fake input tax claims immediately and the pre-reform volume of evasion was huge. The sudden impact means that the reform acted as a big shock to large number of units involved in fraudulent practices. CREST reform only gives an express legal cover to an already existing administrative tool in real time.[27] The denial of input tax is no longer dependent upon costly and time consuming audits which could only cover small fraction of these units and even when these audits materialized in establishing something concrete against one supplier they would have no legal effect for the units down in the chain. But after the reform, a fake supplier cannot easily circumvent the system because, unlike the pre-reform state, the tax authorities can declare its purchases invalid without any need of a formal audit. If this supplier now issues sales invoices to another supplier in next months, then these sales would also be invalid immediately. In this way, the network of fake suppliers collapses because the legal act of invalidating the invoices would already have occurred prior to the purchase by beneficiary units down in the chain. This explains the "sudden" drop in domestic input tax credit claims. It shows that firms were buying inputs from unregistered sectors and were using fake invoices to claim input tax credits to lower their tax liability. The reform had increased the probability of recovery substantially. Assuming that the cost of obtaining a fake invoice remained the same, the specification at 1.10 experienced a jump on the left side of inequality without affecting the right side. A rise in p_2 was enough to make evasion non-feasible.

A natural question arises as to why such action was not taken earlier. But it should be noted here that the success of reform depended on three crucial factors. First, Pakistan had switched to compulsory electronic filing of returns and annexures few years before the reform. Prior to that both manual and electronic filings were allowed.

[27]The concept of real time here means that a firm is legally denied input tax claim as soon as the deadline to file return for a particular month is over.

Data entry of monthly returns and its annexures took a while before they could be fed in the system. Second, the tax machinery had gained the necessary expertise over the years through steady roll out of computerized information solutions. This made the task of whole scale implementation easier by providing the human resource necessary on the ground. Third, Pakistan had up to date IT infrastructure at all tax offices owing to a decade long tax reform program supported by the World Bank.

The implications of these results also extend in two other directions. First, the unregistered sector does not feel the pressure to get registered because their buyers can manage invoices. An indirect implication of the reform is a pressure on informal sector to get registered as they can face stronger competition from the registered sector when they fail to give an invoice to their buyers who can no longer buy the invoice via DMT. Second, the tax machinery could focus on other enforcement tasks instead of fruitless and incomplete audits, investigations and litigation. Obviously, the results of this paper cannot measure these positive spillovers.

Firms differ from each other in many respects. Firm's structure and business type, affect the firm behavior. I analyze effect of the reform based on firm heterogeneity on account of these factors. I divide firms into manufacturers and non-manufacturers by their structure and also study firm behavior by business type (partnership, sole proprietorship or company).

1.5.2 Firm Heterogeneity Analysis

Firms can be further subdivided into different categories based on the structure of firms, nature of their businesses and principal business product they sell. The firms have three options for their business type: i) Sole proprietorship (ii) AOP (Association of Persons or partnership) or a (iii) Company. These categories arise from income tax statutes because three different types are taxed at different brackets. Sole proprietorship is not required to register as a firm and is taxed on the individual's income tax return. AOP has a different income tax rate bracket and companies are taxed at the corporate tax rates. Companies are governed by the Companies Ordinance, 1984 and regulated through Securities and Exchange Commission of Pakistan (SECP). A firm files for VAT registration under one or many of the following categories determined by the nature of its business; 1) Manufacturer (2) Wholesaler (3) Distributor (4) Exporter (5) Importer (6) Retailer (7) Service provider (8) Others. However, registration as manufacturer requires physical visit by tax inspector to verify the address, machinery installed, utilities connections and numbers etc. Although, the law does not bar tax authorities from visiting premises of non-manufacturers but the physical visit as a registration requirement is rare. VAT registration requires additional information regarding nature of the business and general classification of the products which the business shall sell.[28] Therefore, the businesses will also differ according to the goods they manufacture or trade. These heterogeneous characteristics of firms give rise to different possibilities.

Proposition 1: All fake suppliers would be registered as Sole Proprietors.

Proof: Sole proprietorship has the lowest cost of registration because it has fewer documentation and regulatory requirements compared to AOPs and companies. From (1) & (2) above and for a given p_1, p_2, π and l, we have $\theta_{sole} < \theta_{AOP} < \theta_{company}$ which implies that G^a, G^b and G^c are higher and hence from (2), the expected utility would be higher if a sole proprietorship is used as fake invoice supplier. Hence the proof. Intuitively, it also makes sense to use lowest cost registration category to issue fake invoices.

Proposition 2: No manufacturing unit shall be a fake supplier.

Proof: VAT registration as manufacturer requires physical visit and verification by the tax authorities, therefore, for a fake supplier $\theta_{non-manufacturer} << \theta_{manufacturer}$. Similar to the logic used for the proof of proposition 1, supplier of fake invoices would not register as manufacturer.

Proposition 3: The magnitude of fake input tax to the output tax would be higher for larger firms.

Proof: The larger firms already have tax consultants or lawyers on their panel. Therefore the additional legal fee for audit and defending cases in courts determined by l is reduced. By (2), this increases G_b and G_c and the term on the right side of 1.10 also increases which implies the evasion window becomes larger and the maximum expected utility increases more compared to smaller firms who face higher l.

[28] The Act does not place any restriction on selling the goods other than the products mentioned in your registration documents and the information is only meant for statistical purposes but the firms may loose on some incentives if they relate to a particular industry and distributed through VAT return/status. However, the firm can always add or subtract products by simply filing an online request.

(Manufacturer vs. Non-manufacturer): The behavior of the manufacturing firms differs from non-manufacturers. In the context of missing trader fraud, fake suppliers are non-manufacturing units whereas final beneficiary of the fraud is often a manufacturing unit.[29] However, the possibility of using these networks of fake suppliers by non-manufacturing entities cannot be ruled out ab-initio, especially for retailers, service providers, exporters and distributors. Figures 1.5 & 1.6 together with Table 1.3 and Table 1.4 show the balanced and unbalanced panel results for manufacturers and non-manufacturers respectively. The comparison group remains the same i.e. all exporting firms. Coefficients on interaction dummies for four pre-quarters show parallel trend. Domestic input tax claims for both categories drop immediately after the reform. Manufacturers are on average twice the size of a non-manufacturer in the treated group and if the behaviors were identical, the drop in manufacturer's input tax claims should be twice as large compared to the non-manufacturers. But manufacturer's input tax fell by 2.15 million PKR for the balanced panel whereas it dropped by 2.47 million PKR for the non-manufacturers. This result also indirectly supports proposition 2. The higher drop by non-manufacturing units can only be explained by the underlying missing trader fraud as described in preceding sections. Non-manufacturers in the treatment group can have many missing traders who are generating fake input tax so that they can show a corresponding output tax in their returns which can be utilized by their buyers as input tax. Figure A.1 shows a drop in input tax claims for both manufacturers and non-manufacturers. The drop for non-manufacturers based on heterogeneity by business types ranges from 40 to 90 percent with an average effect of approximately 70 percent.

(Heterogeneity by Business Type): Business types allowed in the law are "AOP" (Association of Persons), "Company" (any incorporated entity), "Individual" (Sole proprietorship) and "FTN" (Free Tax Number or Government Agencies) (see Appendix A.1.1 for definitions). FTNs are omitted in the plots as they are special numbers issued to governments such as provincial and local governments for purchasing goods for their own use.[30] Input tax claims drop across each category which shows that fraud is rampant across all business types. Traditional models for developed countries predict that companies being larger organizations would not commit outright frauds. Figure A.2 plots the trends for these business types. In the present

[29] The definition of manufacturer in the law is very broad and thus even a very small processing or repackaging activity etc. changes firm's category to manufacturer.

[30] Government agencies do not have an incentive to evade tax and they are also required to withhold VAT at the time of purchase which ensures deposit of true input tax at the time of purchase.

case, Panel A shows trends for the comparison (or control) group. These firms remain largely unaffected with a small increase in mean logged input tax for companies and a very small drop in domestic input tax claims of sole proprietorships and AOPs. Panel B shows that domestic input claims for companies in the treated group declined significantly.

The estimates of empirical specification at (8) are listed at Tables 1.5, 1.6, 1.7 and coefficients are plotted in figures 1.7 to 1.9. Figures 1.7 and 1.8 plot the evolution of trend for sole proprietorships and partnerships through interaction dummies of four pre and post quarters. In case of sole proprietorships and AOPs the drop is 60-70% compared to their pre-reform levels. These two categories show fairly similar decline in tax credit claims. This result runs counter to proposition 1, which suggested that only sole proprietors would be indulging in missing trader fraud as invoice mills. Massive drop in the claims of sole proprietors partially supports proposition 1. The results do not fully support proposition 1 because the decline in domestic input tax credits of partnerships is also very large and points to existence of invoice mills in these firms. Compared to pre-reform base levels, some categories (distributors and wholesalers) within these firms show drops as high as 90% on average.

Results also support proposition 3 which points to the involvement of bigger firms such as companies in fraudulent behavior. Traditional models for developed countries do not predict fraudulent behavior by companies. But my results show that the same may not be true in developing countries. The incorporated entities were not immune from the market pressure exerted by the evasion across the sectors. Their input claims fell by a massive PKR 4.6 million per entity on average within the first quarter of the reform which is approximately 30% of pre-reform levels. These results support findings in Best et al. (2015) who found that companies paid 70% less tax when they switched from turnover based income tax to profit and loss based income tax. Their results showed that the companies are inflating purchases to reduce profits. My results show that they are also using fake input tax invoices to lower their VAT liability. Together these results, point to a very high evasion in the formal private sector. These companies include many large public sector entities such as the government owned energy sector firms who have no incentive to evade which implies that the overall figure of 30% decline may be very large if the government owned enterprises can be separated.

1.5.3 Robustness of the Results

The results are robust to alternate specifications which can be of concern here. Most importantly, I check whether the results change materially by changing the time period to months instead of quarters and extending the regressions to full eight years of data instead of using one year pre and post reform. Table A.2 lists the results for unbalanced panels by using monthly tax periods. The results are similar to quarterly specifications in Table A.3 and approximately one third of the quarterly values. Table A.3 lists the robustness test of the results at Tables 1.2 to 1.7 for the full time period from the year 2009 to the first quarter of year 2017. These checks show that the results are robust to a very long time period and the change to months from quarters would have no meaningful impact on the results.

My results show that CREST reform curbed DMT fraud. Response of the affected firms is immediate and lasting. The fraud was prevalent across various business categories and types. Real time checking greatly enhanced the chances of recovery which forced this rampant evasion to decline significantly. But this also raises an important question on extent of VAT evasion in developing countries with limited state capacity.

1.6 Conclusion and Extensions

The analysis in this paper shows that VAT in developing countries does not achieve revenue efficiency automatically. Third party reporting in itself does not guarantee that evasion would be curtailed. As long as the enforcement capacity is not high enough, the evasion would be substantial and across the board. In such an environment, VAT imposes tough documentation and compliance requirements on businesses as well as the tax administration without a corresponding benefit. This also lends credence to Emran and Stiglitz (2005) argument that VAT may not be a good choice for developing countries with a limited state capacity.

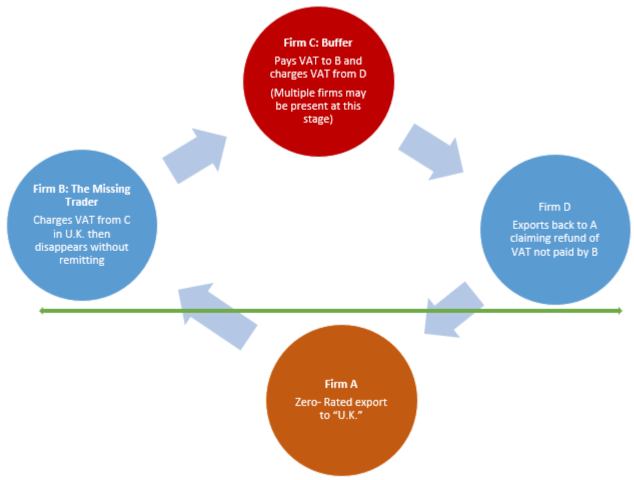
The fact that Pakistan curbed this evasion after nearly two decades of VAT introduction merits greater introspection and should not be considered a natural outcome for other developing countries with VAT regimes. The system used in Pakistan is successful because it is unique in three material respects. First, Pakistani tax administration has been carrying out reforms with significant technical and financial support from the World Bank which greatly enhances its ability to install, operate and execute innovative, latest and advanced computational infrastructure across the tax machinery. Many developing countries greatly lack this institutional expertise.

Second, Pakistan's constitution gives unprecedented powers to Federal Government in tax recovery matters as long as it does not interfere with fundamental human rights. For example, the provisions enacted in Pakistani law can easily be declared unconstitutional in other countries as it practically shifts the burden of proof to the accused. Pakistani courts do provide relief to individual firms in taxation matters but seldom strike down taxation provisions altogether. The constitution also limits the power of stay granted by the superior courts to six months against a tax recovery but not against audit. To the best of my knowledge, this reform and similar other laws were never challenged in any court except for individual relief. Third, Pakistan also operates an extensive zero rating and exemption regime which greatly reduces the workload on tax machinery from refunds and audits thereby freeing resources and reducing compliance costs, in absence of which, these reforms cannot be implemented successfully.

The results are also important in the context of European Carousel Fraud because a system of real time checking at the domestic level can guard against the MTIC fraud without involving support from other countries or other innovations such as origin based taxation, CVAT and VIVAT. As developed nations, EU countries can easily replicate Pakistani model subject to the constitutional limitations discussed above.

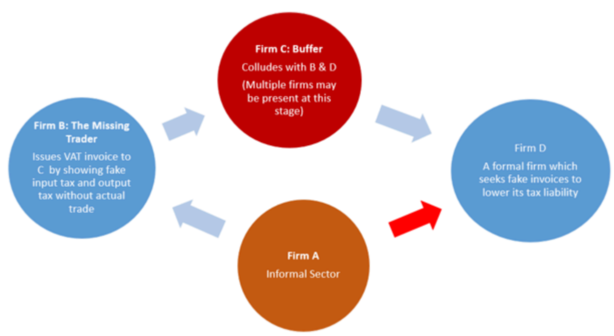
These results support the theoretical predictions and provide impetus for an extended analysis of input tax evasion. The emphasis in the tax debate on VAT should not only relate to sales suppression but also to purchase inflation and plugging the loopholes which result in revenue leakage at large scale. This paper does not cover sectoral, geographical and network based dynamics of the missing trader fraud which needs further exploration.

Figure 1.1: Missing Trader Fraud



A- European Carousel Fraud

Explanation: This figure is adapted from Keen & Smith (2006). The European carousel fraud operates in a chain and principally relies on one or more suppliers going "missing" after issuing invoices without depositing the tax collected. For example, firm A exports goods to U.K. from an EU member country and gets the VAT refund of tax paid on its purchase. Firm B imports the goods without paying any tax at import and supplies it to firm C (multiple firms can exist at this level). Firm B now goes missing without remitting the tax to government but is invoice issued to firm C is a valid instrument to claim input tax credit. Firm D buys the goods from firm C and exports it back to the same member country and claims refund on strength of the invoice issued by C. U.K. government ends up paying the amount which was never deposited in the exchequer.

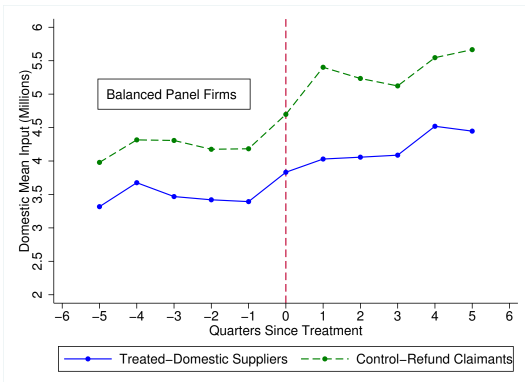


B- Domestic Missing Trader Fraud

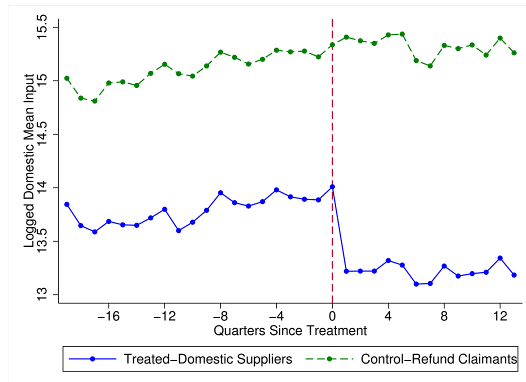
Explanation: In domestic missing trader fraud, Firm D is actually purchasing goods from A but would not be able to claim input tax because firm A is not registered to issue VAT invoice. Firm D, A or both, now collude with suppliers of fake invoices who have set up fake units that only generate invoices without conducting actual business. These units are registered in different geographical jurisdictions making audit difficult for the revenue authorities. They also provide plausible deniability for firm D which operates in formal sector. In case of Pakistan, the different territorial jurisdictions of the tax offices provide an incentive to operate a missing trader fraud as the audit and enforcement can only be conducted by the office having geographical jurisdiction and the invoices of missing traders can be difficult to verify.

Figure 1.2: Raw Data- Domestic Input Tax Credit

(a) Raw Data-Balanced Panel Only



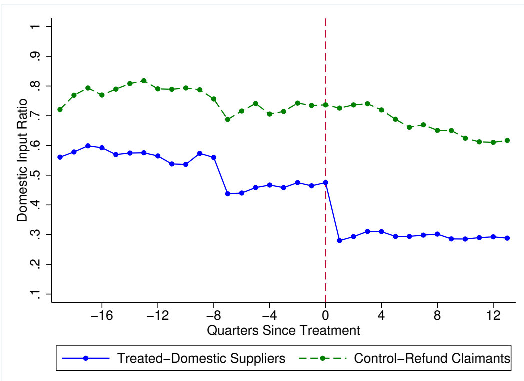
(b) Raw Data- Quarterly Input Tax



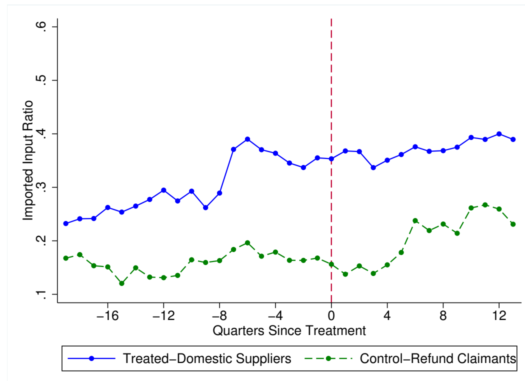
Explanation: 2(a) The graph plots the quarterly mean input tax of the firms in balanced panel for the period from Jan-March 2012 to June 2014 such that Lead 5 is the quarter Jan-March 2012 and Lag 5 is the quarter July-September 2014. The mean for comparison group in this graph has been scaled down by a factor of three such that 4 million point corresponds to 12 million for comparison group. 2(b) The graph shows parallel trend by plotting logged mean quarterly domestic input tax credit of control and treated groups in PKR millions. The reform occurs at dashed vertical line (quarter April-June 2013) which is then used as a reference to show lead and lag quarter time periods. The drop is sudden and the treated group again follows the control group but with a bigger mean difference giving support to the identification strategy especially with reference to common shock assumption. Decline is approximately 70 log points or 50 percent.

Figure 1.3: Domestic and Import Input Credits' Ratio to Total Credits

(a) Raw Data-Ratio of Domestic Input Credit to Total Credit



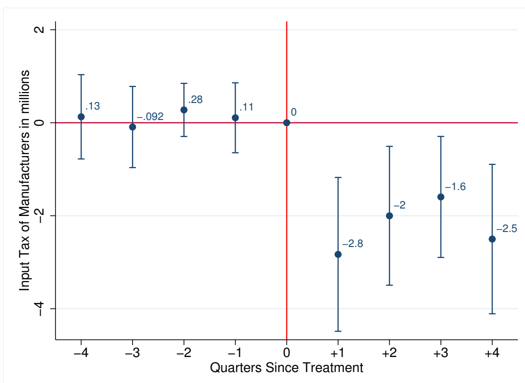
(b) Raw Data- Ratio of Import Input Credit to Total Credit



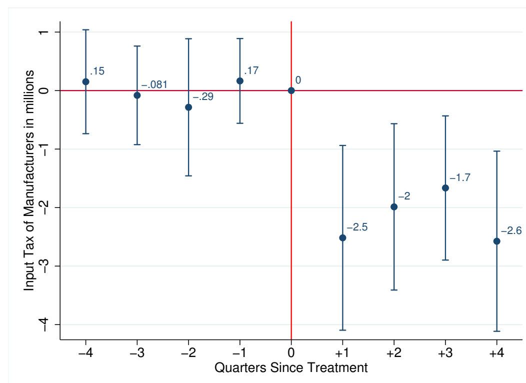
Explanation: The reform occurs at dashed vertical line which is then used as a reference to show lead and lag quarter time periods. (Panel A) The graph shows pre-reform parallel trend by plotting the ratio of mean quarterly domestic input tax credit to total input claimed by both comparison and treated groups. Sudden post-reform drop for the treatment group shows that domestic credit claims declined as a percentage of total claims by approximately 50%. (Panel B)The graph of ratio of imported input tax credit to total input tax shows that the imports remained stable for both groups which provides evidence that the trend is not driven by a reduction in business or other factors which should normally affect purchases in overall terms both domestic as well as imported.

Figure 1.4: Aggregate Effect of CREST on Domestic Input Tax Claims

(a) Balanced Panel



(b) Unbalanced Panel

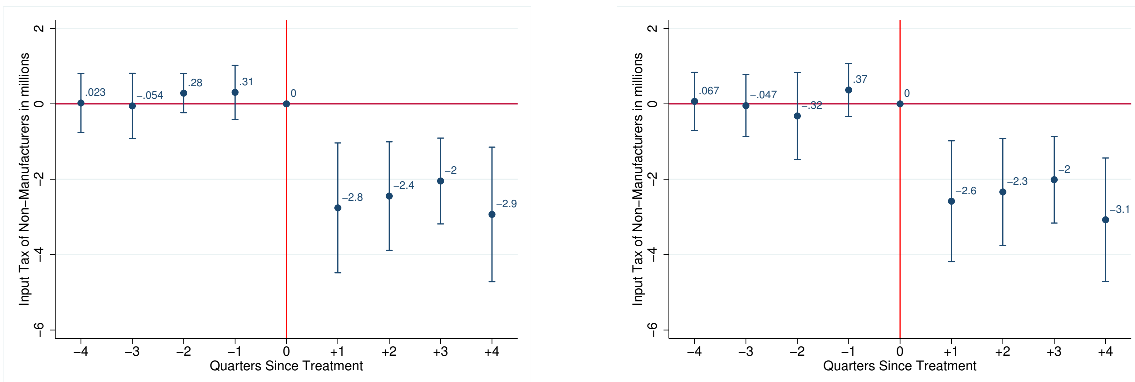


Explanation: The figure plots point estimates of DD dummies for quarter specific interactions with 95% confidence level to rule out any pre trend (for details see Table 1.2). The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The regression covers the period from April 2012 to June 2014 such that Lead 4 is the quarter April -June 2012 and Lag 4 is the quarter April -June 2014. Panels A & B show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level.

Figure 1.5: Effect on Domestic Input Tax Claims of Manufacturers

(a) Balanced Panel

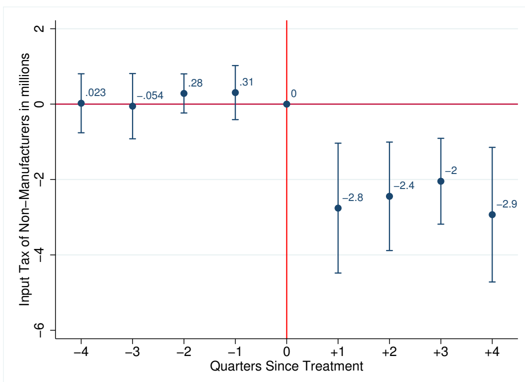
(b) Unbalanced Panel



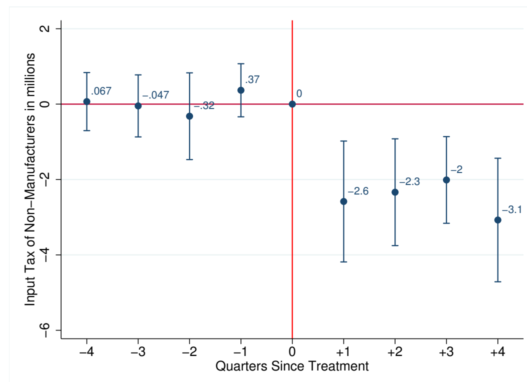
Explanation: The figure plots point estimates of DD dummies for quarter specific interactions with 95% confidence level to rule out any pre trend (for details see Table 1.3). The comparison group is same as for previous figure but only the manufacturers from the treatment group are included. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The regression covers the period from April 2012 to June 2014 such that Lead 4 is the quarter April-June 2012 and Lag 4 is the quarter April-June 2014. Panels A & B show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level.

Figure 1.6: Effect on Domestic Input Tax Claims of Non-Manufacturers

(a) Balanced Panel



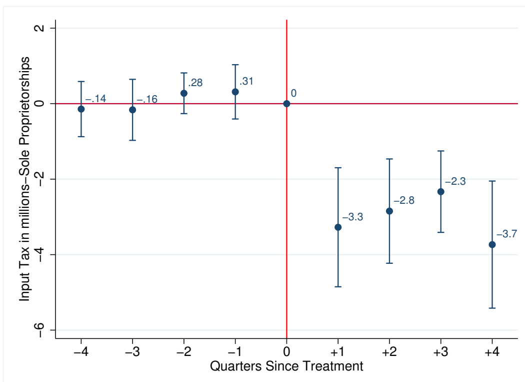
(b) Unbalanced Panel



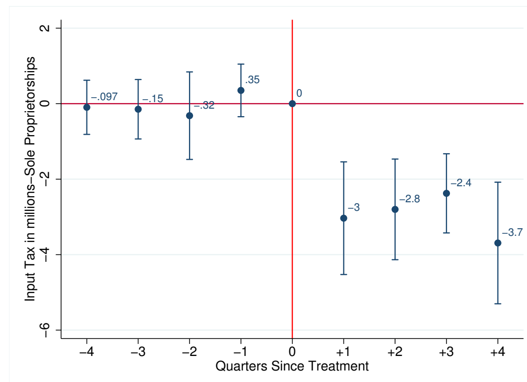
Explanation: The figure plots point estimates of DD dummies for quarter specific interactions with 95% confidence level to rule out any pre trend (for details see Table 1.4). The comparison group is same as for previous figure but only the non-manufacturers from the treatment group are included. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The regression covers the period from April 2012 to June 2014 such that Lead 4 is the quarter April -June 2012 and Lag 4 is the quarter April -June 2014. Panels A & B show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level.

Figure 1.7: Effect on Domestic Input Tax Claims of Sole Proprietorships

(a) Balanced Panel



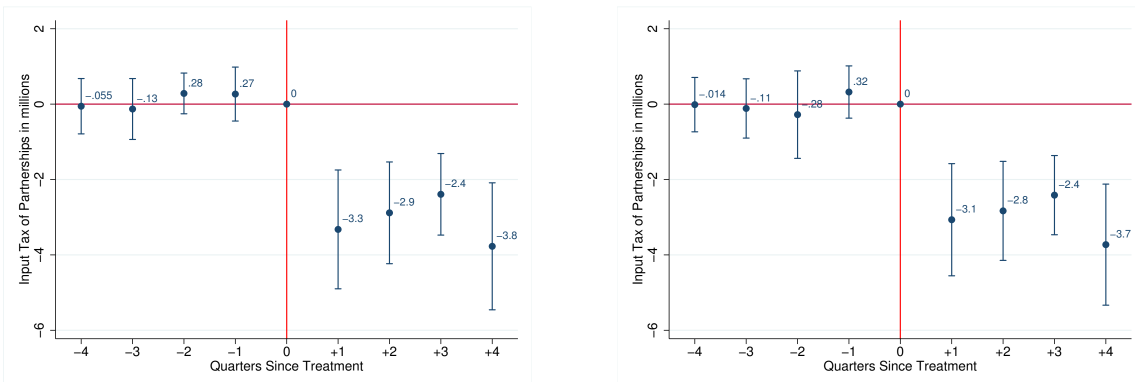
(b) Unbalanced Panel



Explanation: The figure plots point estimates of DD dummies for quarter specific interactions with 95% confidence level to rule out any pre trend (for details see Table 1.6). The comparison group is same as for previous figure but only the sole proprietorships (businesses owned by only one individual) from the treatment group are included. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The regression covers the period from April 2012 to June 2014 such that Lead 4 is the quarter April -June 2012 and Lag 4 is the quarter April -June 2014. Panels A & B show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level.

Figure 1.8: Effect on Domestic Input Tax Claims of Partnerships

(a) Balanced Panel (b) Unbalanced Panel

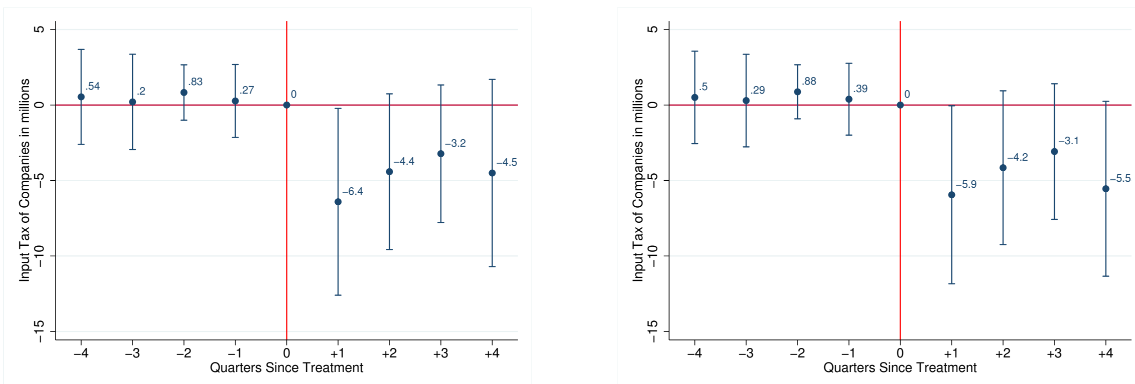


Explanation: The figure plots point estimates of DD dummies for quarter specific interactions with 95% confidence level to rule out any pre trend (for details see Table 1.5). The comparison group is same as for previous figure but only the firms registered as partnerships in the treatment group are included. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The regression covers the period from April 2012 to June 2014 such that Lead 4 is the quarter April -June 2012 and Lag 4 is the quarter April -June 2014. Panels A & B show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level.

Figure 1.9: Effect on Domestic Input Tax Claims of Companies

(a) Balanced Panel

(b) Unbalanced Panel



Explanation: The figure plots point estimates of DD dummies for quarter specific interactions with 95% confidence level to rule out any pre trend (for details see Table 1.7). The comparison and treatment groups both have the firms registered as companies only. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The regression covers the period from April 2012 to June 2014 such that Lead 4 is the quarter April-June 2012 and Lag 4 is the quarter April -June 2014. Panels A & B show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level.

Table 1.1: Descriptive Statistics

	Domestic Suppliers	Exporters	Others
	(Treatment)	(Comparison)	
All Firms			
Domestic Input Tax (Mean)	706,928	4,093,938	5
Std. Deviation	36,900,000	68,800,000	126
# Observations	6,214,612	626,090	2,617,535
Manufacturers			
Domestic Input Tax (Mean)	1,140,941	3,003,698	7
Std. Deviation	32,400,000	55,600,000	162
# Observations	1,791,292	546,030	411,623
Non-Manufacturers			
Domestic Input Tax (Mean)	531,167	11,500,000	5
Std. Deviation	38,600,000	126,000,000	118
# Observations	4,423,320	80,060	2,205,912
Companies			
Domestic Input Tax (Mean)	4,825,110	12,300,000	5
Std. Deviation	110,000,000	122,000,000	137
# Observations	679,688	197,840	231,599
Partnerships			
Domestic Input Tax (Mean)	217,788	230,911	5
Std. Deviation	1,943,833	934,920	124
# Observations	1,156,853	199,296	498,439
Sole Proprietorships			
Domestic Input Tax (Mean)	183,904	329,014	5
Std. Deviation	1,458,975	8,198,891	125
# Observations	4,376,500	228,615	1,887,235
Government Agencies			
Domestic Input Tax (Mean)	35,700,000	1,880,429	1
Std. Deviation	376,000,000	2,633,667	0
# Observations	1,607	228,615	262

Notes: Domestic input tax figures are in Pakistani Rupees (100 PKR =1 $). The comparison firms are the ones who had claimed refund in excess of 1 million PKR for the five year period before the reform, making all the remaining firms treatment group except "Others". "Others" column shows the firms who had very little or no input tax credit for the five year period before the reform (less than 10,000 PKR in total) and therefore, remain out of the purview of analysis for this paper.

Table 1.2: Revenue Impact of CREST Reform

Domestic Input Tax (PKR in Millions)				
	(1)	(2)	(3)	(4)
	Balanced	Balanced	Unbalanced	Unbalanced
DD (Post June 13 × Domestic Input Tax)	-2.36***		-2.22***	
	(0.66)		(0.64)	
Lead 4		0.12		0.15
		(0.45)		(0.44)
Lead 3		-0.06		-0.05
		(0.44)		(0.42)
Lead 2		0.29		-0.29
		(0.26)		(0.59)
Lead 1		0.24		0.29
		(0.37)		(0.36)
Lag 1		-2.80***		-2.53***
		(0.86)		(0.81)
Lag 2		-2.26***		-2.19***
		(0.73)		(0.7)
Lag 3		-1.90***		-1.91***
		(0.6)		(0.59)
Lag 4		-2.72***		-2.86***
		(0.82)		(0.77)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes
Number of Groups	43,928	43,928	115,669	116,038
N	395,352	438,539	670,213	717,469

Notes: Table displays the main coefficients as well as coefficients on quarter specific interaction dummies for firm level regressions. Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. Leads and lags variables are DD dummies for quarter specific interactions to rule out any pre trend (for plot see figure 1.4). The regression covers the period from March 2012 to June 2014 such that Lead 4 is the quarter March-June 2012 and Lag 4 is the quarter March-June 2014. Column (1) (2), and (3) (4) show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level and shown in parenthesis. See Table A.3 for robustness checks. *** denotes that the results are significant at 1% level.

Table 1.3: Revenue Impact on Manufacturers

Domestic Input Tax (PKR in Millions)				
	(1)	(2)	(3)	(4)
	Balanced	Balanced	Unbalanced	Unbalanced
DD (Post June 13 × Domestic Input Tax)	-2.15***		-2.02***	
	(0.67)		(0.65)	
Lead 4		0.13		0.15
		(0.46)		(0.45)
Lead 3		-0.09		-0.08
		(0.45)		(0.43)
Lead 2		0.28		-0.29
		(0.29)		(0.60)
Lead 1		0.11		0.17
		(0.38)		(0.37)
Lag 1		-2.83***		-2.52***
		(0.84)		(0.81)
Lag 2		-2.00***		-1.99***
		(0.76)		(0.73)
Lag 3		-1.60***		-1.66***
		(0.66)		(0.63)
Lag 4		-2.5***		-2.58***
		(0.82)		(0.79)
Input Tax on Imports	-0.01	0.04	-0.01	0.04
	(0.05)	(0.05)	(0.05)	(0.05)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes
Number of Groups	21,323	21,323	33,374	33,484
N	191,907	212,937	241,149	263,597

Notes: Table displays the main coefficients as well as coefficients on quarter specific interaction dummies for firm level regressions in case of manufacturers. Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. Leads and lags variables are DD dummies for quarter specific interactions to rule out any pre trend (for plot see figure 1.5). The regression covers the period from March 2012 to June 2014 such that Lead 4 is the quarter March-June 2012 and Lag 4 is the quarter March-June 2014. Column (1) (2), and (3) (4) show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level and shown in parenthesis. See Table A.3 for robustness checks. *** denotes that the results are significant at 1% level.

Table 1.4: Revenue Impact on Non-Manufacturers

Domestic Input Tax (PKR in Millions)				
	(1)	(2)	(3)	(4)
	Balanced	Balanced	Unbalanced	Unbalanced
DD (Post June 13 × Domestic Input Tax)	-2.47***		-2.34***	
	(0.66)		(0.64)	
Lead 4		0.02		0.07
		(0.4)		(0.39)
Lead 3		-0.05		-0.05
		(0.44)		(0.42)
Lead 2		0.28		-0.32
		(0.27)		(0.59)
Lead 1		0.31		0.37
		(0.37)		(0.36)
Lag 1		-2.76***		-2.58***
		(0.88)		(0.82)
Lag 2		-2.45***		-2.34***
		(0.73)		(0.72)
Lag 3		-2.05***		-2.01***
		(0.58)		(0.59)
Lag 4		-2.93***		-3.07***
		(0.91)		(0.84)
Input Tax on Imports	-0.25	-0.21	-0.17	-0.15
	(0.38)	(0.4)	(0.32)	(0.28)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes
Number of Groups	28,971	28,971	89,700	89,972
N	260,739	289,173	491,079	522,346

Notes: Table displays the main coefficients as well as coefficients on quarter specific interaction dummies for firm level regressions in case of non-manufacturers. Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. Leads and lags variables are DD dummies for quarter specific interactions to rule out any pre trend (for plot see figure 1.6). The regression covers the period from March 2012 to June 2014 such that Lead 4 is the quarter March-June 2012 and Lag 4 is the quarter March-June 2014. Column (1) (2), and (3) (4) show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level and shown in parenthesis. See Table A.3 for robustness checks. *** denotes that the results are significant at 1% level.

Table 1.5: Revenue Impact on Partnerships

Domestic Input Tax (PKR in Millions)				
	(1)	(2)	(3)	(4)
	Balanced	Balanced	Unbalanced	Unbalanced
DD (Post June 13 × Domestic Input Tax)	-2.94***		-2.78***	
	(0.61)		(0.59)	
Lead 4		-0.06		-0.01
		(0.37)		(0.37)
Lead 3		-0.13		-0.11
		(0.41)		(0.4)
Lead 2		0.28		-0.28
		(0.28)		(0.59)
Lead 1		0.27		0.32
		(0.36)		(0.35)
Lag 1		-3.32***		-3.07***
		(0.8)		(0.76)
Lag 2		-2.88***		-2.83***
		(0.69)		(0.67)
Lag 3		-2.39***		-2.41***
		(0.55)		(0.54)
Lag 4		-3.77***		-3.73***
		(0.86)		(0.82)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes
Number of Groups	13,496	13,496	27,879	27,928
N	121,464	134,775	177,296	191,492

Notes: Table displays the main coefficients as well as coefficients on quarter specific interaction dummies for firm level regressions in case of partnerships. Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. Leads and lags variables are DD dummies for quarter specific interactions to rule out any pre trend (for plot see figure 1.8). The regression covers the period from March 2012 to June 2014 such that Lead 4 is the quarter March-June 2012 and Lag 4 is the quarter March-June 2014. Column (1) (2), and (3) (4) show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level and shown in parenthesis. See Table A.3 for robustness checks. *** denotes that the results are significant at 1% level.

Table 1.6: Revenue Impact on Sole Proprietorships

Domestic Input Tax (PKR in Millions)				
	(1)	(2)	(3)	(4)
	Balanced	Balanced	Unbalanced	Unbalanced
DD (Post June 13 × Domestic Input Tax)	-2.88***		-2.72***	
	(0.62)		(0.6)	
Lead 4		-0.14		-0.1
		(0.37)		(0.37)
Lead 3		-0.16		-0.15
		(0.41)		(0.4)
Lead 2		0.28		-0.32
		(0.28)		(0.59)
Lead 1		0.31		0.35
		(0.37)		(0.36)
Lag 1		-3.27***		-3.03***
		(0.8)		(0.76)
Lag 2		-2.85***		-2.80***
		(0.7)		(0.68)
Lag 3		-2.33***		-2.38***
		(0.55)		(0.54)
Lag 4		-3.73***		-3.69***
		(0.86)		(0.82)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes
Number of Groups	32,106	32,106	84,558	84,867
N	288,954	320,489	489,989	524,581

Notes: Table displays the main coefficients as well as coefficients on quarter specific interaction dummies for firm level regressions in case of sole proprietorships. Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. Leads and lags variables are DD dummies for quarter specific interactions to rule out any pre trend (for plot see figure 1.7). The regression covers the period from March 2012 to June 2014 such that Lead 4 is the quarter March-June 2012 and Lag 4 is the quarter March-June 2014. Column (1) (2), and (3) (4) show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level and shown in parenthesis. See Table A.3 for robustness checks. *** denotes that the results are significant at 1% level.

Table 1.7: Revenue Impact on Companies

Domestic Input Tax (PKR in Millions)				
	(1)	(2)	(3)	(4)
	Balanced	Balanced	Unbalanced	Unbalanced
DD (Post June 13 × Domestic Input Tax)	-4.59*		-4.68**	
	(2.4)		(2.35)	
Lead 4		0.54		0.5
		(1.6)		(1.56)
Lead 3		0.20		0.29
		(1.61)		(1.56)
Lead 2		0.83		0.88
		(0.93)		(0.91)
Lead 1		0.27		0.39
		(1.23)		(1.21)
Lag 1		-6.41**		-5.95**
		(3.15)		(3.00)
Lag 2		-4.42		-4.15
		(2.63)		(2.60)
Lag 3		-3.22		-3.08
		(2.32)		(2.29)
Lag 4		-4.5		-5.55*
		(0.86)		(2.95)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes
Number of Groups	6,670	6,670	12,754	12,780
N	60,030	66,607	83,365	90,274

Notes: Table displays the main coefficients as well as coefficients on quarter specific interaction dummies for firm level regressions in case of companies. Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. Leads and lags variables are DD dummies for quarter specific interactions to rule out any pre trend (for plot see figure 1.9). The regression covers the period from March 2012 to June 2014 such that Lead 4 is the quarter March-June 2012 and Lag 4 is the quarter March-June 2014. Column (1) (2), and (3) (4) show the results for a balanced and unbalanced panel respectively. Standard errors are clustered at firm level and shown in parenthesis. See Table A.3 for robustness checks. *** denotes that the results are significant at 1% level.

Chapter 2

Is Minimum the Maximum? Tax Burden on Informal Sector in VAT: Evidence from Pakistan

Developing countries with large informal sectors collect a very low percentage of their GDP as tax revenue. Many businesses remain out of the tax net despite having large volume of sales and significant profits. These countries, therefore, rely on tax instruments that can place an adequate burden of taxes on informal firms. This desire to tax informal sector is a major reason why indirect taxes, considered regressive by economists, still bring major share of tax revenues through import tariffs, excises and other consumption taxes. Value added tax (VAT) remains the most widely used indirect tax on consumption across the world including developing countries.
A critical debate in public finance deals with value added tax (VAT) and its welfare effect in developing countries who have large informal sectors. These countries have lowered their import tariffs in response to WTO agreement and greatly reduced excise duties over the last few decades. Reducing the magnitude and scope of tariffs and excises without a compensating tax instrument is revenue inefficient and also relieves informal sectors from indirect burden of taxes. Consequently, the countries enacted VAT regimes to make up for the revenue lost and increase tax burden on informal sectors of their economies. The underlying premise is that VAT would generate sufficient revenue by taxing the imported goods and their subsequent supply chain. At the same time, informal sector would bear the tax burden because staying out of registered supply chain would result in denial of tax credit to an informal firm which is available to a formal one. However, on the other hand, if firms can evade tax easily it would greatly undermine both revenue efficiency and burdening of informal sector hypotheses. A revenue neutral VAT in these circumstances would be welfare decreasing not increasing (Emran and Stiglitz 2005).

The welfare increasing argument relies on the assumption that production efficiency of formal sector and indirect burden of VAT on informal sector makes it a better choice compared to trade taxes (Keen 2008). Apparently, informal sector would be burdened through VAT to the extent it pays tax on purchases obtained from formal sector. However, for government, revenue gained from informal sector consumption may actually be very small if importers can suppress their sales value by declaring a very low or no value addition. It implies that the tax paid at import is now lower because of lower tariffs and anticipated compensation through tax on post-import value addition cannot be collected. Additionally, the informal sector which buys these imported goods bears lesser burden of tax. Therefore, presence of large evasion undermines welfare increasing argument and strengthens the opposite view.

In developed countries, informal sectors are very small and the importers sell goods to other formal firms who demand a VAT invoice to claim credit for the tax already paid. In developing countries, importers sell goods to both formal and informal sector buyers. The formal buyer would need a true value invoice because they can claim credit only with a verifiable VAT invoice. But informal buyer has no need for an invoice and may even prefer not getting one because it would help her conceal her business and remain out of the ambit of tax authority. In such cases, importer can easily declare little to no value addition[1] and evade VAT. Tax authorities tackle this value addition suppression by imposing minimum values addition thresholds to calculate output tax due on sales.[2] It implies that tax liability of an importer is flat below this threshold. This minimum tax threshold introduces a non-linear kink in an otherwise linear tax schedule. This kink provides an incentive to firms with higher value additions to bunch at or below the kink in a manner similar to kinks in non-linear income tax schedule. If the firms were reporting true value addition they should remain unaffected by introduction of this kink. If the firms bunch, then this behavior indicates evasion by suppression of sales to informal sector. Recent evidence points to a strong incentive to evade in environments where tax authority does not have the advantage of in-built information trail to obtain third party information such as the retail sales to the consumers (Naritomi 2019). Similarly, self-regulating mechanism of VAT can collapse if buyers and sellers both have an incentive to report lower sales value. Because buyer is an unregistered person or informal firm, she gains by lower declared price as long as she shares some of the benefits of these lower value declarations. The VAT structure built on the concept of arm's length transactions between buyers and sellers starts breaking down under these conditions.

This chapter exploits the kinks created by minimum value addition thresholds

[1] Value addition is simply the difference of gross sales and gross purchases. In case of an importer value addition would, thus, simply be a difference of sales value and value assessed at import by customs authorities.

[2] Minimum taxes in corporate and personal income taxes are very common.

in Pakistani VAT regime to calculate volume of evasion on account of value addition suppression by non-manufacturer importers (hereinafter referred simply as importers). Pakistani law requires importers to pay a minimum value addition tax, based on an arbitrary single value addition threshold for all firms irrespective of the type of goods being imported. If subsequent actual value addition of firms is higher than this presumptive value addition, they deposit excess differential tax liability with their VAT returns. But in case their actual value addition is less than the presumptive value addition, they cannot claim adjustment of excess tax paid at the time of import. Therefore, firms have no incentive to cheat if their value addition is less than arbitrarily set threshold but a strong incentive to bunch below the threshold when their actual value addition is higher than threshold.

My setting is unique because it can easily disentangle income effect from VAT evasion benefit. The VAT evading firms earn dual benefits because they not only lower their VAT liability but also use this evasion to decrease their income tax liability. Presence of this additional income effect prevents an accurate estimation of VAT evasion under normal settings. It becomes difficult to estimate whether this misdeclaration is driven solely by VAT evasion benefits or not. The present setting is unique because income tax liability of an importer is fixed based on its import value and not its value addition or actual profits. They would get no benefit from declaring a lower sales value nor would they be paying anything extra no matter how large a sales value they declare. Consequently, the whole evasion can be solely attributed to VAT benefit. The estimates in this chapter are therefore more reliable than any other estimates in VAT literature because they, unlike other estimates, disentangle the effect of income tax enforcement and evasion from VAT enforcement and evasion.

I use administrative data for the universe of monthly VAT returns filed in Pakistan for tax years 2009 to 2016 to estimate evasion using value addition thresholds. Using the variation created by different minimum thresholds in different years for the non-manufacturing firms exclusively engaged in imports, I can estimate an average measure of evasion at post importation stage. I build on the methodology developed to estimate elasticity of taxable income for non-linear tax schedule by Saez (2010) and Chetty et al. (2011). I find high bunching with approximately 50% of all firms bunching on average around the threshold. It translates to an evasion rate of 78%. This large bunching can only be explained by negligible adjustment costs and the fact that firms far above the threshold can easily bunch around minimum value addition threshold. In personal income taxes where individuals loose wages by working less, the incentive to bunch is limited to a very small window where tax savings compensate for the loss of income. In present case bunching results from misreporting and firms are not loosing any actual sales. They are only declaring wrong sales to tax authorities. Hence, given the incentive, huge bunching is surprising but possible.

This chapter adds to three strands of literature. First, it adds to the debate on revenue efficiency and enforcement of VAT in developing countries with large informal sectors. This literature provides evidence of VAT evasion in developing countries, see (Pomeranz 2015; D. R. Agrawal and Zimmerman 2019; Naritomi 2019; Waseem 2019; Waseem 2020). However, this literature is exclusively restricted to manipulation of purchases and sales either at local supplies or export. I add to this debate by providing evidence of value addition manipulation by importers. Another important debate in VAT evasion literature centers around impact of VAT withheld at import stage on informal sector with two opposite views. Emran and Stiglitz (2005) argue that high import tariffs and excise duties mean that large informal sectors share the burden of indirect taxes adequately. When VAT replaces these high import tariffs, informal sector's tax burden gets lowered and, therefore, a revenue neutral VAT would decrease welfare. Keen (2008) opposes this view by arguing that VAT is collected at import stage and then importers also charge tax on subsequent value addition thereby imposing sufficient burden on informal sector who cannot claim any credit of tax paid at import stage. I provide, to the best of my knowledge, first empirical evidence regarding this debate. I estimate widespread evasion post importation which implies that the argument of Emran and Stiglitz (2005) is valid regarding VAT implementation in developing countries. Third, a growing literature in personal and corporate income taxes estimates evasion response at a kink by bunching estimator (for example (Blomquist et al. 2021; Gelber, Jones, and Sacks 2020; Card et al. 2015)). Henrik Jacobsen Kleven (2016) explains the methodology of bunching and provides a detailed review. Almunia and Lopez-Rodriguez (2018) use bunching methodology for enforcement notches in Spain. Liu et al. (2021) use bunching to estimate response at VAT notches created by registration thresholds in VAT. However, bunching estimators have not been used to estimate responses at VAT kinks. Sow and Gebresilasse (2020) use bunching at VAT registration notch in Ethiopia to model firm behavior at VAT notches. I provide evidence of bunching at VAT kinks created by minimum taxes and use the program developed by Chetty et al. (2011) to estimate bunching at VAT kinks.

2.1 Institutional Setting

The federal VAT in Pakistan operates mainly under standard VAT regime where firms deduct input tax paid on their local and imported purchases from the tax due on sales. They either deposit the difference with their VAT return or claim refund of excess paid tax. Pakistan also has large informal sectors which remain outside formal tax regime. Presence of these large informal sectors often forces the government to enact special procedures to guard against revenue leakage through sales to these informal sectors of economy. Importers who do not have manufacturing facility sell their imported goods at wholesale markets which largely remain informal. Informal buyer has no concern with sales value reported by their sellers because truthful reporting of seller has no marginal benefit for buyer. On the other hand, misreporting of true sales value provides a direct benefit to seller who can keep differential amount of tax due as profit. Therefore, these importers have a strong incentive to deviate from truthful reporting and declare lower than actual value addition. They collect price inclusive of tax but can keep a portion of this collected tax by declaring a price lower than the actual price. For revenue authorities, cost of audit and establishing a case of sales suppression without any third party information is very high. Audit observation is very difficult to withstand judicial scrutiny because the goods are of foreign origin and open market price is rarely available independently.[3] Lower sales declaration also implies lesser income tax liability and this increases the incentive for evasion.

In an effort to check the double incentive described above, Pakistani tax regime introduced two different deviations from standard VAT and income tax regimes. The importers have to pay a fixed percentage over and above their assessed value at import stage as a "final" tax on their income tax liabilities. This means that these firms can neither adjust the income tax paid at the time of import nor the income tax department would demand any amount over and above the withholding tax collected at import stage. This effectively rules out any income effect on account of lesser sales declaration. Therefore, the setting provides an interesting scenario where all the effect can be attributed to VAT evasion.

Minimum value addition (MVA) regime operates under a special procedure. The law stipulates additional tax to be paid over and above the standard tax rate. This additional tax varied from 2% to 3% from the year 2008 to 2016. At a VAT rate of

[3]This sale value suppression does not take into account under invoicing at import stage and assumes that assessed customs value of the goods on which this tax was withheld is fair. If the customs value of the goods is lower than the market value, it would be another source of evasion but is not the subject of this research.

16%, this translates to 12.5% and 18.75% value addition respectively. For example, if an importer imported goods whose custom duty inclusive value is assessed at 100 PKR then she would pay a 16% tax amounting to 16 PKR and an additional 2 PKR for the 2% MVA tax. This implies that the government is assuming that the importer would sell these goods which cost him 100 at a value of 112.5 (112.5*0.16=18). The threshold is therefore sensitive to MVA tax rate as well as the standard tax rate. Threshold is calculated using a simple relation between after-import sales value $V_s$and import value V_I, given by

$$V_s = V_I \left(1 + \frac{t_m}{t}\right) \tag{2.1}$$

where t is the standard tax rate and $t_m$is the minimum tax rate at imports. During this period, the tax rate also changed once from 16% to 17% implying the threshold also shifted slightly from 18.75% to 17.65% (see Figure 2.1). All these changes were passed with the beginning of new fiscal years, therefore, I can aggregate monthly data to yearly data to determine behavior through bunching more precisely.

2.1.1 MVA on Non-Manufacturer Importer

The importers (who are selling goods imported without further processing) in Pakistan operate under a “minimum value addition” (MVA) regime. The rationale for imposing an MVA is to ensure that sectors prone to evasion on supply of goods (on account of supplies made to informal sector) are paying their due share. However, if the actual value of supply is more than this minimum then importer shall pay excess tax due at the time of filing return. The minimum VAT collected at the time of import is not refundable.[4] Additionally, the importers falling under this MVA regime are exempt from Audit which implies that an importer has no enforcement pressure to declare a value addition more than the one required by MVA threshold. Only three factors can induce her to declare more sales than the statutory minimum, i) Arm’s length transaction between registered buyers and sellers which requires truthful reporting, (ii) maintaining double books of accounts without being traced and (iii) lack of information about the loophole that prevents the firm from evasion.[5] These factors can induce adjustment frictions similar to those observed by Chetty et al. (2011)

[4]In Pakistan, for the period under consideration, the importers were taxed on income under a “Final Tax” regime wherein a fixed percentage of import value (inclusive of all taxes) is considered as the income tax liability and collected at the time of import.

[5]Although a third response driven by an altruistic responsibility to the society which demands no tax evasion is possible but I shall ignore it in this chapter in line with tax evasion literature (See (Allingham and Sandmo 1972)). Also, in present case, its not possible to differentiate between a response arising from lack of information and altruism.

which prevent income tax payers from reducing their income.

2.2 Theory

Consider a firm which imports goods at a value V_I and sells them at a value $V_s$such that $V_s = V_I + V_a$where $V_a$is the value added after import. If C is the cost of getting this value addition then profit

$$\prod = V_I + V_a - C$$

When a tax at rate τ is imposed then

$$V_s * \tau = (V_I + V_a) * \tau$$

is the revenue remitted to the government by the importer. In this ideal scenario, VAT would have no impact on the profit of a firm and would be an economically neutral tax. However, when government imposes a minimum value addition requirement this economic neutrality vanishes which would reduce economic efficiency but may or may not improve revenue efficiency.

Now I consider an importer who is operating in an economy with a large informal sector. In this case, many buyers of the imported goods may not be registered firms. Therefore, even in the case of arm's length transactions, these buyers do not derive any benefit out of correct reporting of sales value by their sellers.$V_I$is the only reference value for the tax authority because any previous transactions of goods occurred outside importing country.[6]

The importer can misreport sales value to increase her profit through evasion. If $V_g$is the value reported to government such that $V_g \leq V_a$ and ϕ is the adjustment cost function then profit is given by

$$\prod = V_I + V_a - C - \underbrace{(V_I + V_m) * \tau}_{\text{Tax paid at import}} - \underbrace{(V_g - V_m) * \tau}_{\text{Tax due with return}} + \underbrace{(V_a - V_g) * \tau - \phi(\omega)}_{\text{Evasion Benefit}} \quad (2.2)$$

I define parameter $\omega\epsilon\{0,1\}$ as ratio of sales to unregistered buyers to total sales such that $\phi(\omega)$ is increasing in ω and concave. An additional restriction on ϕ is that $\phi \rightarrow 0$ as $V_g \rightarrow V_a$which is evident from the nature of the problem that if there is no under declaration then there would be no adjustment costs as well. In a standard

[6]$V_I$is the customs value which according to WTO rules is the transaction value between importer and exporter subject to certain additions and restrictions.

VAT there is no limit on$V_g$and an importer can claim refund or adjustment of tax paid at import stage by declaring a negative value addition post importation. In Pakistani case, government restricts any adjustment or refund of excess VAT paid by importer through imposing a minimum value addition (V_m) requirement and collects that portion of tax at the time of import. The importer pays an additional tax with the return when $V_m \leq V_a$. It imposes an additional constraint on$V_g$such that $V_m \leq V_g \leq V_s$. It also implies that evasion would only be feasible if $V_a > V_m$ and the evasion expression is positive. In cases where $V_m \leq V_a$, the firm's adjustment cost function would determine the most feasible V_g. I now examine firms in different scenarios and their expected response in light of equation (2.2).

When Actual Value Addition is less than MVA; This is the case when $V_a \leq V_m$. Here, firm declares correct value to the government so that $V_a = V_g$and there is no evasion by the firm as positive term for evasion response goes to zero. Tax due with return is now negative or refundable but the government has expressly prohibited adjustment or refund of this excess paid amount. Therefore, this is an additional cost to a firm whose actual value addition is lower than the minimum threshold.

When Actual Value Addition is more than MVA; In this case, a firm can evade taxes and add an amount over and above its profit absent an MVA requirement, because it is now out of purview of any audit or enforcement measure and only restriction it faces is determined by ω which is lower for a lower volume of sales to unregistered buyers. It charges the market price inclusive of tax from unregistered buyer in informal sector but remits a lesser amount to government thereby increasing its profits. This firm would try to declare a value addition as close to the threshold as possible. These firms locate at or around V_m and that would produce a corresponding hole above this threshold.

Consider firms with preferences defined over price received at true value minus the tax paid and the effort given by actual value addition percentage such that the utility function is given by $U(V - T(V), \frac{V}{n})$, where V is the value added, $T(V)$ is output tax function and n is the ability of the firm. The heterogeneity in ability can be described by a density distribution $f(n)$. I assume that the ability distribution, the tax system and preferences are smooth. Also, the optimization by firms produces a smooth value addition distribution. The kinked tax function can now be constructed based on the fact that below the minimum value threshold there is no change in output tax liability. This implies that

$$T(V) = \begin{cases} V_m * \tau & V \leq V_m \\ (V_g - V_m) * \tau & V > V_m \end{cases}$$

If there's no kink then firms would simply show a value addition at which they can maximize their utility and would locate along the $(1-\tau)$ slope line. However, when kink is introduced, they have an incentive to bunch around the kink. The firms can reasonably defend their value addition at or close to threshold determined by the government. The adjustment cost of locating at the kink reduces considerably for the firms whose actual value addition is higher because of audit exemption. Nonetheless, they would still need to keep two different books of accounts and could be charged with fraud if this manipulation is caught. An implication of this new scenario is that the bunching around threshold is a very feasible and adjustable strategy . The slope of resulting line is 0 before the kink and $1-\tau$ after the kink. Figure 2.2 shows the case of a firm whose actual value addition is dV higher than the threshold V_m. This firm would be better off declaring a value addition less than its actual value addition. The present problem is analogous to traditional theory of bunching at kinks. The only difference here is a much sharper kink which after allowing for optimization frictions and adjustment costs is similar in functional form to the bunching methodology employed by Chetty et al. (2011).

2.2.1 Implications of a Kinked Output Tax Function

Introduction of minimum tax threshold creates a kinked output tax liability function as described above. For any firm this functions presents a problem only when its actual value addition is less than V_m. In such case this firm would be forced to pay an amount equal to $(V_m - V_a) * \tau$ over and above its true liability where V_a is actual value of sales. Also, this firm would have no incentive to misreport its true sales value. For firms with value additions greater than V_m, there is an incentive to report as low a value to government V_g as possible. A lower V_g implies lower tax payment than the tax due and higher profit. This firm can face three possible scenarios:

i) All Supplies to Formal Firms: If all buyers are from formal sector they would demand a true value invoice to claim input tax credit already paid on the goods and firm would report $V_g = V_a$. The government would loose no revenue.

i) All Supplies to Informal Firms: If all buyers are informal firms, importer would have an incentive to declare $V_g = V_m$. In this case the firm would bunch very close to threshold because the adjustment cost is small and revenue authority has no meaningful way of knowing true sales value.

i) Supplies to Both Formal and Informal Firms: If the firm is selling goods to both formal and informal sector then it faces adjustment costs because it would report formal sector sales correctly as in case (i) above but would suppress sale value for sales made to informal firms as is the case at (ii) above. This firm may or may not be able to bunch near threshold depending upon her particular circumstances.

2.3 Identification Strategy

MVA in Pakistani case is selected arbitrarily by government and is same irrespective of the goods being sold. The imported goods can range from agricultural produce to sophisticated engineering products with hundred of different countries of origins. In this case, it is reasonable to assume that actual value addition for the underlying distribution of value addition post importation would be smooth in absence of this threshold. If Ψ is this underlying distribution then $\Psi = \int_{-\infty}^{+\infty} d\Psi$. The real response should show no consistent bunching at any arbitrarily determined point. But given the incentive available for importers to bunch around the kink, an evasion response is expected which would produce an excess mass at or below the kink and a hole above it. I approach this problem in the same manner as developed by Saez (2010) with minor modifications as described below.

Let $H_0(V)$ be the cumulative frequency distribution of value addition when there's a constant marginal tax rate τ_0 and $h_0(v) = H_0'(V)$ be the corresponding density distribution. I assume that sales value are smoothly distributed according to the distribution $h(v)$. The heterogeneity in profits (including profit through evasion) depends on two factors- inherent entrepreneurial ability of the firm represented by sales value and its preference to evasion- both captured by a utility function defined on two parameters $U(\pi, v)$. In a pre-kink world with evasion, the slope of profit after value addition tax line is $(1 - \tau)$ .When a minimum value addition threshold is introduced, then slope of corresponding line is zero because the line is now flat till the point V_m. Introduction of kink at this point means that firms who were previously showing values below the MVA point remain unaffected but firms whose value additions are above this level would be better off by bunching below this minimum point as shown in figure 2.2. I assume that there is an infinitesimally small segment below $V_m$where slope is $(1 - \tau + d\tau)$ so that the evasion (e) by declaring lesser value addition than actual is captured by the relationship

$$\frac{dV_m}{V_m} = e\frac{d\tau}{1-\tau} \tag{2.3}$$

In a heterogeneous population of firms there would be now a new distribution of firms $\Psi(v, e)$ such that $h(V_m) = \int_e \Psi(V_m, e)de$ and average evasion at this value addition level V_m is then given by $\bar{e} = \int_e e\Psi(V_m, e)de/h(V_m)$. The presence of a kink at

V$_m$implies that the number of firms bunching at V_m is $dB = \int_e eV_m(\frac{d\tau}{1-\tau})\Psi(V_m, e)de = \bar{e}h(V_m)V_m(\frac{d\tau}{1-\tau})$.

To account for kink points which are not small enough, I consider a parameterized model in which evasion response is assumed to be a quasi-linear and constant across firms, so that the utility function has the form

$$U(\pi, v) = \pi - \frac{n}{1+\frac{1}{e}}(\frac{v}{n})^{1+\frac{1}{e}}$$

where n is entrepreneurial ability parameter of the firms which has a cumulative frequency distribution $F(n)$, a corresponding density given by $f(n)$ and normalized to one.[7] The bunching would therefore identify evasion because I am assuming that affinity to evade is similar across firms (although their actual response may be attenuated by adjustment cost). I maximize $U(\pi, v)$ subject to the constraint $\pi = V - T(V) = (1-\tau) * v$. The resulting first order condition: $1 - \tau - (\frac{v}{n})^e = 0$ can be rewritten as:

$$v = n(1-\tau)^e$$

If there was no tax ($\tau = 0$) then simply $v = n$ which means I can interpret (n)as declared value addition in absence of VAT. For the constant marginal tax rate case with τ_0,$v = n(1-\tau_0)^e$and therefore $H_0(V) = Pr(n(1-\tau_0)^e \leq v) = F(v/(1-\tau_0)^e)$ which means that corresponding density function is $h_0(v) = f(v/(1-\tau_0)^e)/(1-\tau_0)^e$. Now introduce a convex kink by increasing the marginal tax rate to $\tau_1$such that $\tau_1 > \tau_0$ when value addition is above V_m. Introduction of a kink above means that for $v > V_m$, $v = n(1-\tau_1)^e$, therefore, in a corresponding manner as done earlier cumulative frequency distribution and density function are $H_1(V) = Pr(n(1-\tau_1)^e \leq v) = F(v/(1-\tau_1)^e)$ and $h_1(v) = f(v/(1-\tau_1)^e)/(1-\tau_1)^e = h_0(v((1-\tau_0)/(1-\tau_1))^e) \times ((1-\tau_0)/(1-\tau_1))^e$. If I denote $h(V_m)^+$ and $h(V_m)^-$ as the right and left band around V_m, then $h(V_m)^- = h_0(V_m)$ and $h(V_m)^+ = h_0(V_m((1-\tau_0)/(1-\tau_1))^e) \times ((1-\tau_0)/(1-\tau_1))^e)$.

I denote the density of realized value addition and its cumulative distribution by $h(v)$ and $H(v)$so that for the actual value addition below the kink at V_m, $v = n(1-\tau_0)^e$ for $n < V_m/(1-t_o)^e$. This implies that $h(v) = h_0(v)$ for $v < V_m$. Because $v = n(1-\tau_1)^e$ for $n > V_m/(1-t_1)^e$therefore, it follows that for $v > V_m$, $H(v) = Pr(n(1-\tau_1)^e \leq v) = F(v/(1-\tau_1)^e)$ and $h(v) = f(v/(1-\tau_1)^e)/(1-\tau_1)^e = h_0(v((1-\tau_0)/(1-\tau_1))^e) \times ((1-\tau_0)/(1-\tau_1))^e$. When $v \to V_m$ then the left and right limits of

[7]The quasi-linearity assumption may not be necessary in this particular case as income tax is already fixed and collected at the time import as final liability but would be required for cases where income tax is not collected under final tax type mechanism applied here.

the function are simply $h(V_{m-}) = h_0(V_m)$ and $h(V_{m+}) = h_0(V_m((1-\tau_0)/(1-\tau_1))^e) \times ((1-\tau_0)/(1-\tau_1))^e$ respectively. The firms with value additions higher than $V_m$over an interval $V_m + \Delta V_m$would now bunch at or below the kink point such that*

$$\frac{\Delta V_m}{V_m} = \left(\frac{1-\tau_0}{1-\tau_1}\right)^e - 1 \tag{2.4}$$

This equation is similar to the relationship developed by Saez (2010) for bunching in the piece-wise linear income tax schedule with VAT minimum threshold point V_m replacing the bunch point $z^*$for income threshold. It is also a general form of equation (2.3). Therefore, the bunching of firms represents a trapezoid over and above the counterfactual distribution given by $h_0(v)$which can be evaluated from trapezoidal rule of approximation for a definite integral by using the relationship $h(v) = h_0(v)$ and equation (2.4) as:

$$B = \int\limits_{V_m}^{V_m+\Delta V_m} h_0(v)dv \backsimeq \Delta V_m \frac{h_0(V_m) + h_0(V_m + \Delta V_m)}{2} = \Delta V_m \frac{h(V_{m-}) + {}^{h(V_{m+})}/\left(\frac{1-\tau_0}{1-\tau_1}\right)^e}{2} \tag{2.5}$$

By combining equations (2.4) and (2.5), I get the following relationship;

$$B = V_m \left[\left(\frac{1-\tau_0}{1-\tau_1}\right)^e - 1\right] \frac{h(V_{m-}) + {}^{h(V_{m+})}/\left(\frac{1-\tau_0}{1-\tau_1}\right)^e}{2} \tag{2.6}$$

V_m is directly observable and τ_0 is simply $V_I * \tau_m$, knowing B, $h(V_{m-})$ and $h(V_{m+})$ the evasion response can be estimated. This would be similar to the estimation used by Saez (2010). Estimation of this equation assumes that there is no friction. Also, this would need an approximate value of V_I for all firms or at least an average measure of V_I. Therefore, it would be better to employ the method proposed by Chetty et al. (2011) which relies on estimating counterfactual density. Because of my large sample, I can estimate a counterfactual density and excess bunching mass taking into account the adjustment costs and frictions.

There are two important differences that must be taken into account here. First, income response of an individual together with net of marginal tax rate can identify elasticity of taxable income but in case of VAT I also need a measure of volume of sales to calculate the volume of evasion. Ideally, I can assume that sales value of bunching and non-bunching firms are same on average but this assumption would need actual support from sales and bunching behavior of the firms in different deciles or quartiles.

If, there is sufficient variation between sales volume of two categories then I would need to incorporate a measure for average sales of bunching firms. Second, unlike income tax, firms cannot evade by actually manipulating their purchase and sales as the customs has already recorded purchases. Also, output goods are in same state as they were imported without undergoing any manufacturing process. Therefore, there can be no actual response and all the variation is due to manipulation of sales value. For registered buyers, seller cannot manipulate invoice value because buyer would need that invoice to claim credit of tax already paid. For unregistered buyer, however, there is no such restriction and seller can easily report any value. But this misreporting would, at least, require maintaining two different books of accounts.

2.4 Data and Empirical Estimation

I use universe of data of monthly returns filed by VAT firms from the tax year 2008 to 2016 in Pakistan. The return data captures over 100 different return columns which is required to disentangle different types of sales, purchases, carry forwards etc. [8] I aggregate data by year because all changes in threshold were introduced from the beginning of a financial year. Descriptive statistics are shown in Table 1. Manufacturing firms must have a manufacturing facility of their own and do not fall under MVA regime (For a definition of important terms see Appendix).

Pakistani MVA regime for importers deducts tax on a presumptive value addition. Before 4th June 2011, the importers were assessed at 2% of the import value (inclusive of all import stage taxes) as an MVA based tax and paid at the time of Customs clearance. Using relation given by equation (2.4), it translates to an MVA of 12.5% and implies that a firm can show any value addition in its return but it would not matter in terms of tax liability if its less than 12.5%. But if it exceeds this threshold then excess tax would be deposited with the return. Additionally, importers falling under this category were exempt from audit till year 2012 but they were also exempt in subsequent periods as this category was excluded from audit program. [9] From tax year 2012, this tax was increased to 3% or 18.75% MVA. Because tax rate post MVA threshold is linear and uniform, it implies that any excess bunching at or around threshold is arising because of suppression of sales value of the imported goods and If the importers were not directly or indirectly involved in evasion then there should

[8] Some firms are only required to file a quarterly return but aggregation by year automatically takes care of this issue.

[9] The statutory exemption from audit was removed through Notification No. S.RO. 592(I)/2012, dated 1st June, 2012, w.e.f. 2nd June, 2012, reported as PTCL 2013 St. 673.

be no change post reform in their actual value addition and sales declared to formal and informal sectors. Figure 2.1 shows the changes in threshold over eight years from tax year 2009 to 2016. Empirical estimation of parameters through bunching requires calculation of counterfactual density. I follow STATA code of Chetty et al. (2011) to estimate counterfactual density, bunching mass b and number of firms bunching near the kink.

2.4.1 Graphical Evidence

First step in bunching estimation is the graphical evidence. Figure 2.3 plots the histograms for each year from 2009 to 2016. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 2 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 6 percent would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. It is evident from these plots that there is significant bunching near the threshold point in each year. Figure 2.3 shows yearly pattern of declared value addition and density around threshold points is very high. I start these histograms from a value addition of negative 20 percent. On the left side, the number of firms gets steadily lower until it almost reaches zero or negligible value. On the right side of threshold, I extend these histograms till 80% value addition. Naturally, there would be firms declaring more value addition than 80% or lesser value addition than -20% but extending analysis beyond these points on either side is unlikely to impact any estimation significantly. Figure 2.4 shows distribution of firms for different thresholds. By combining data of year 2009-2011, I plot distribution for 12.5% threshold (Panel A). Similarly, by combining data for years when threshold was same, I plot distribution for other two thresholds (Panels B & C). Panel D shows a cumulative picture with different thresholds shown by red lines. Plots for threshold periods are very similar to the yearly plots. These plots show massive bunching compared to the bunching literature in income taxes. The main reason for this high bunching are two. First, firms do not loose any real earnings by reporting less sales instead they are evading taxes and adding to their profits. Second, unlike registration thresholds where a firm may also gain by declaring its true size and formalizing, these importing firms are already formal and are not producing new products by employing labor. They have no similar incentive to declare sales correctly.

The fact that firms almost seamlessly change their behavior in 2012 when threshold increases to 18.75% suggests that these bunching firms were coming from the right side of bunch point because otherwise they would have no incentive to bunch at higher tax liability point. Another important point is that when threshold changes to 18.75%, firms shift to new point. It suggests that these firms are not showing much friction in adjusting their response. It further strengthens the argument that bunching is coming from right side of distribution. For example, if a firm had actual value addition of 40% but was declaring only 12.5% value addition it was very easy for this firm to show 18.75% value addition as the adjustment cost, however little it may be, would go down as well.

2.4.2 Estimating Counterfactual Density

The next step is to estimate a counterfactual density. An inherent problem with this setting is that there is no prior period available when this MVA threshold did not exist. Had that been the case, I could take distribution of firms prior to the introduction of kink and measure the changes. This threshold is general and applicable to all firms dealing in any products as long as they are not manufacturers. Normally, there would be a natural case for using manufacturers as the comparison group and use their distribution as comparison. But manufacturers generally import raw materials or intermediate goods and then further process them to sell a final product. Their value additions are difficult to correlate with their imports. Many manufacturers use locally procured inputs and they do not pay fixed and final income tax at import stage. Also, locally procured inputs of manufacturers can be exempt goods as well as taxable goods. It can confound value addition estimates for manufacturers. Therefore manufacturers cannot be a good counterfactual in this particular case.

Another option is to assume a uniform distribution over a range of value addition. The case for uniform distribution is supported by very flat and small mass outside a reasonably wide bunching window centered on the bunch point. I can select bunching windows based on graphical distribution. For example, in figure 2.3 for year 2009, there is very little mass below zero and only a slightly thicker mass from zero to 5 percent. Similarly, on right side of bunch point after 35% bin there is very small mass till 40% bin and then it starts becoming negligible. Years 2010 and 2011 are fairly similar but from 2012 onward both left and right sides of distribution become thicker and more gradual. This would make comparison across different years and threshold periods difficult.

I, therefore, use the method developed by Chetty et al. (2011), to estimate a counterfactual density. In figure 2.5 I plot histogram for all importing firms in Pakistan. I calculate value addition as difference between sales value and import value reported in their returns. I divide it by import value to generate percentage value addition. I divide firms in bins with each bin representing one percent. I then plot bin counts in each bin to construct a histogram. Red vertical line shows the bin of threshold which in year 2009 is 13% bin. The figure shows that there is a sharp spike around MVA threshold. Now the underlying assumption of this analysis is that absent this MVA threshold, distribution of firms by their value addition percentage would be smooth. It implies that counterfactual distribution would be different from this histogram. The excess bunching mass, b is proportional to evasion. But this bunching mass is diffuse over a bunching window where the firms having value addition higher than threshold are locating around threshold point to avoid higher VAT payments. To measure this mass, I fit a polynomial to the points plotted in figure 2.5 by leaving out the data in area around bunch point. The polynomial is estimated by following regression:

$$C_j = \sum_{i=0}^{m} \beta_i^0.(V_j)^i + \sum_{i=-R}^{R} \gamma_i^0.\mathbf{1}\{V_j = i) + \varepsilon_j^0 \tag{2.7}$$

where C_j is the counter for number of firms in bin j, $V_j$is the percentage value addition bin, m denotes the degree of fitted polynomial, $-R$ corresponds to the width of bunching window on left side of threshold and R shows the bunching window on right side of bunch point. Chetty et al. (2011) describe the estimation procedure in detail. I use the same code, therefore, estimation assumptions and calculations are identical in this case. Only difference is that they estimate it over income bins whereas I am estimating it over value addition bins. I estimate a value for parameter b, which is the excess mass in terms of number of firms which locate around threshold kink relative to counterfactual density of value addition distribution. I use a seventh degree polynomial and a window of 12 bins centered on the bunch point. A value of $b = 1$ means that excess mass around kink is 100% of the average height of counterfactual distribution within 6% (R=6) of the kink. I select 6% window because the reform in 2012 shifted MVA threshold by 6% or 6 bins. This 6% window means that when estimating b for second threshold period, there is no overlap for bins on the left (right) with left (right) bins of previous threshold. Using the null hypothesis of $b = 0$, I find that alternate is true in all cases discussed in this chapter at p-value of less than 1%. The standard errors are calculated using bootstrap method with 200 or more iterations. The estimation for all parameters and all estimates of counterfactual shown in tables and figures converge within 10 iterations.

2.4.3 Bunching Response by Year

A key feature of MVA reforms over time is that they were implemented at the start of a financial year. It enables me to aggregate monthly and quarterly returns to get yearly figures for sales and imports. Annual figures substantially reduce noise because I can also see carry forward (brought forward) stocks for each year. Spillover from previous year is recorded in first monthly return for next year. It means that spillover amounts, if any, are very small compared to annual aggregates of imports and sales. Another advantage for yearly aggregation is that I can see sales by quartiles and deciles to see whether smaller importers behave differently from bigger importers.[10]

Figure 2.5 shows yearly estimates of bunching. Panels a, b & c for years 2009,2010 and 2011 plot bunching estimate for the period when threshold was 12.5%. Red vertical line shows the bin of threshold. Bunching is very sharp with bunching mass b, averaging around 18. It means the excess mass is 18 times the height of counterfactual distribution. Number of firms bunching near kink is astonishingly high. Out of a total of 5355 firms, 62% of firms are bunching near MVA threshold in year 2009. The estimates for years 2010 and 2011 are also similar. Given the fact that threshold is arbitrary and same across all products and countries of origin, it is unlikely to be a real response or truthful declaration. Moreover, if the response is real then firms should not shift to higher value addition when threshold changes. However, when this MVA threshold changes to 18.75% in year 2012, firms start bunching at new threshold point. Had the response been real and driven by market considerations only, firms had no need to bunch at new threshold when this threshold is higher than their actual value addition. On the other hand, firms facing reduced adjustment cost in terms of possible lesser penalties and lesser misreporting related adjustments have an incentive to declare closer to declare value addition closer to new threshold. New MVA threshold is higher by 50% in year 2012 but number of firms bunching near threshold is only 6-7% less than year 2011 (Panel d). A movement of bunch point to 6.75% on right side has translated to roughly equivalent drop in percentage of firms bunching near threshold. It implies that actual mass of firms on the left side of previous bunching window was only approximately 7% (or 1% per bin).

Estimates of yearly bunching response are given in Table 2.1. Response of firms is inelastic in terms of value addition and they bunch at MVA threshold as long as threshold does not change but they quickly shift in almost perfectly elastic way to new threshold. It also implies that frictions and adjustment costs are really low.

[10] Imports and sales can also be seasonal which would be problematic for division by quartiles and deciles.

Given this scenario, it is reasonable to predict that if this threshold is removed then firms will declare little to no value addition. In extreme cases, they may start filing for refund paid at import stage by claiming losses at sales stage. In such case, the optimal strategy for revenue authority is to impose a backstop to disallow refunds. In present case, FBR went a step ahead and imposed a minimum tax on these firms. Absent, this MVA tax only 30-40% firms would have declared true value addition.

2.4.4 Bunching Behavior by Period

I also examine response of firms by particular threshold periods. Figure 2.6 plots the distribution of firms in a particular threshold period. First period is from 2009-2011, second from 2012-2013 and third is post 2013. Panel A shows that firms are following their yearly behavior for a particular threshold period. In Table 2.1, value of b are 16.92, 19.13 and 18.52 for years 2009, 2010 and 2011 respectively. The results for period wise response are tabulated at Table 2.2. Value of b for Period 1 is 18.26 which is identical to the yearly average of 18.19 we get from Table 2.1. Similarly for Period 2 and Period 3, values of b are 14.13 and 11.82 from Table 2.2 and corresponding yearly average calculated from Table 2.1 are 14.17 and 11.23. It shows that, over the course of a period, the bunching response of firms is identical with little to no significant deviation from their yearly response. There is a drop of roughly 6% when threshold increase by 6.75%. It shows that there were only few firms whose actual value addition was between two thresholds (12.5% and 18.75%). It implies that only 6% of the firms in previous bunching window were declaring their actual value addition. It leaves me with 55.5% of firms who were bunching near the kink on 12.5% MVA threshold but have now shifted to new threshold. This response can only be attributed to evasion on account of lesser value addition. These results provide strong support and evidence that firms are shifting away from one value addition to another costlessly although they are 50% apart in terms of value added. The numbers of firms bunching near kink is 50% on average across three periods. However, there are some minor yet important deviations in Period 3.

In Period 3, threshold shifts slightly from 18.75% to 17.65% or one bin towards left in bunching estimation figures. Number of firms bunching near kink drops by 5% from Period 2 and bunching becomes more diffuse. Although, these numbers are very small and do not impact the findings stated above, but I need to point a few things to alleviate any concern about this minor drop. The response is driven by more diffused bunching rather than the number of firms bunching below or only a bin above. This change in threshold was only 1 bin movement to the left and it was on account of change in standard tax rate. It is expected that few firms may not be aware of this change because it was not a result of direct amendment in special rules governing these importers. Another factor that might be playing a role here is introduction of 1% further tax on supplies made to unregistered informal buyers in 2014. This additional turnover tax increases adjustment cost for bunching below or around threshold in two ways. First, importer now needs to collect an extra 1% from unregistered buyer, keep track of that amount and deposit it with monthly returns. Second, all else being equal, there is a comparative advantage in selling goods to registered buyer compared to unregistered buyer.

My initial assumption is that there is no significant differences across deciles and quartiles among bunching and non-bunching firms. If bunching behavior is different, then I would need to calculate an average value of sales for bunching and non-bunching firms. I would also need an assumption regarding the actual value addition of a bunching firm. For example, I do not know what is the exact location towards right of bunch point for a particular firm. In income tax bunching, the bunching individual is not very far from bunch point and a mid point assumption is sufficient. The same logic does not work here for two reasons. First, a firm much farther away from bunch point can bunch below threshold. Counterfactual density in all plots does not follow actual distribution till 30 to 40 bins towards right of bunch point. In that case, it would be problematic to assume that an average buncher has actual value addition of say 40% or 50% because the incentive to bunch is increasing with more value addition. In Income tax schedule, economic incentive to bunch around threshold decreases as the taxpayer's income increases beyond bunch point. In this case, reverse is true. Greater the value addition of a firm, more incentive they have to bunch below threshold because bunching is nearly costless. Second, income in itself defines incentive in an income tax scenario. However, I need sales along with assumption on value addition to arrive at volume of evasion. Sales volume and value addition are correlated in a way. Normally, businesses with high turnover or sales can sustain on lower value addition compared to those with smaller sales volumes because of lower average fixed costs. Additionally, value addition changes considerably from nature of product and industry. Finished products can be sold at much higher value addition factor compared to raw materials. I do not have data to see what were the goods actually imported. Therefore, I check heterogeneity based on quartiles to spot any significant differences.

2.4.5 Heterogeneity by Volume of Sales

I analyze response by quartiles to determine heterogeneity on account of sales volume. I construct quartiles and deciles based on yearly sales volume and then group these yearly quartiles and deciles by period to calculate bunching estimates. I do not show results for deciles in plots because they are very similar to their quartile without much variation within quartile. Therefore, I restrict my analysis to quartiles only. Key point for this analysis is whether firms with larger sales volume are bunching more precisely or not. Larger firms face less average fixed cost and that can inflate bunching. On the other hand, larger firms face more cost in terms of double book keeping such as more transactions through banking channels.

Figure 2.3 plots response by quartile for third period. Response for first two periods are described in Figure B.1.1 and Figure B.1.2. Firms in first and second quartile are bunching very precisely. There is very high mass around bunch point in these quartiles. In third and fourth quartile, bunching is more diffused around threshold with heavy left and right tails. However, total number of firms bunching on left of bunch point is similar. It translates to lesser value of b as we go to higher quartiles. But if number of firms who are bunching below threshold does not change, it implies that effect on actual evasion is insignificant. Table 2.3 shows results for third period. Bunching mass b at 20.4 is very high for first quartile and it decreases to only 5.4 for fourth quartile. However, percentage of firms bunching near bunch point is 39% in fourth quartile compared to 56% for first quartile. Higher volume of sales makes precise bunching more difficult but it does not appreciably change behavior in terms of actual evasion as the firms are now less precise but still bunching below threshold.

2.4.6 Elasticity of VAT at Post-import stage

The similarities between bunching window and quartile wise response in first two periods allow me to calculate elasticity of tax collection with respect to changes in tax rate. Change in tax collection in VAT is given by product of (change in number of firms bunching)*(change in width of bunching window of value addition variable)* (average sales) * (tax rate). Figure 2.6 shows that width of bunching window above threshold is same across two periods. Similarly, quartile wise response of first two periods is also identical which implies that average sales of firms in question do not change. Because tax rate is also same in these two periods, therefore, percentage change in number of firms bunching gives the percentage change in tax remitted.

Percentage change in tax rate is given by the percentage difference of minimum tax rate. Therefore, elasticity of tax remitted can be calculate using firms bunching near kink from Table 2.2. The value of elasticity comes to 0.22. Changing threshold by one percent increases tax collection by only 0.22% instead of at least 1% that should be expected absent any evasion. Absent a minimum tax, 78% of the tax would be evaded and only 22% deposited into treasury.

2.4.7 Robustness of Results

I check the behavior by deciles to check whether the results in top and bottom quartiles are driven by very small or very large entities but do not find any significant deviation. I also check whether this behavior is driven by geographic proximity but that also does not exist. The response is not highly concerted at a particular tax office or region.

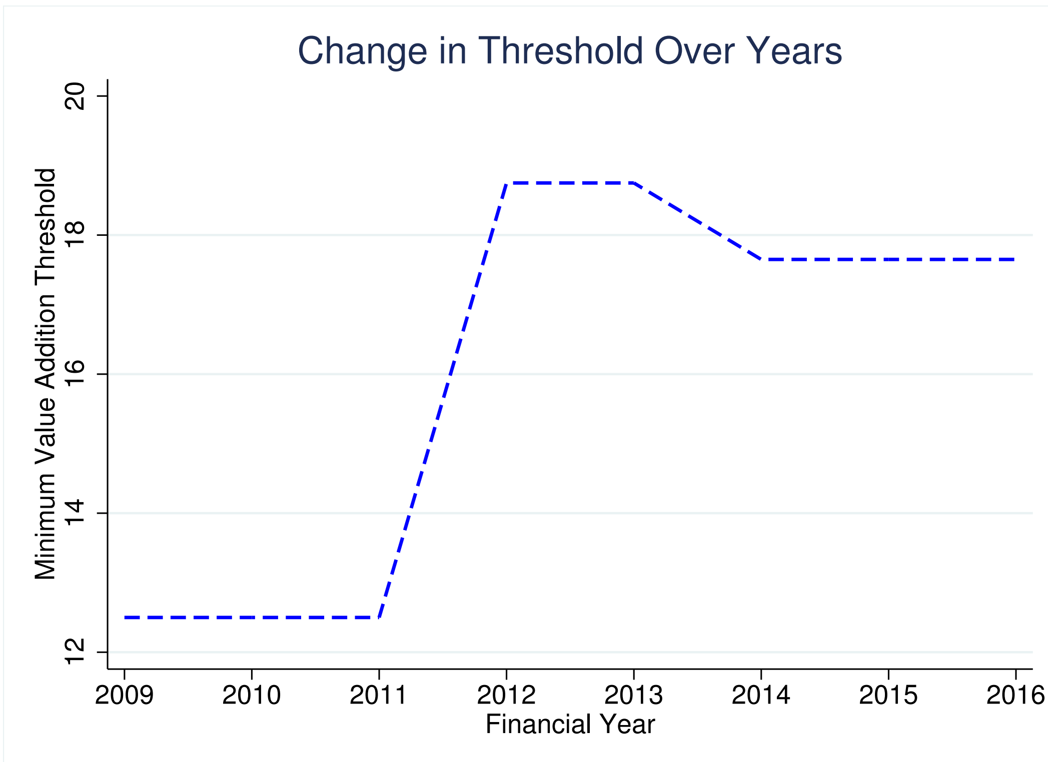
2.5 Conclusion

My results indicate that importing firms are evading VAT and volume of this evasion is approximately 78% with on average 50% of firms bunching at MVA kinks. It means that if MVA threshold is not applied, government would loose significant revenue. This behavior is largely consistent across quartiles by sales volume. The response might be a deviation from economic neutrality concept of VAT but it is revenue efficient. Increasing threshold by 50% automatically increases revenue by 50% but it does not impact many higher value addition type firms adversely as only 6% of firms were actually bearing increased cost of deviation from this economic neutrality. The increased revenue is coming from truthful firms who are now burdened with extra payments. My results provide evidence of significant evasion on account of supplies to informal sector. VAT, in such case, is no longer a double edged sword which rewards registered firms and penalizes unregistered sector. Absent any MVA, importers would under-declare value of supplies to informal sector and pocket the burden intended to be passed on informal sector as profit.

My evidence strengthens the argument of Emran and Stiglitz (2005) that replacing high tariffs with VAT would decrease welfare. It does not find support for counter argument advanced by Keen (2008). In tax literature, there are not many studies who look exclusively at imports in VAT setting. This essay contributes to literature on taxation at import stage and its downstream consequences. Owing to limitations of my data which does not include product specific imports, I am unable to examine commodity wise behavior which may shed more light on behavior of importing firms. Also, I have assumed assessed customs value as reference point. If these import values of firms were higher and they declared less at time of customs clearance then my results are underestimating evasion. However, if these import values were actually lower and customs assessed them higher than transaction value, then my results are overestimating evasion. Because studies and reports on Pakistan's imports show considerable aggregate less declaration at import compared to export value reported at port of shipment, it is reasonable to assume that commercial importers are declaring less than their actual transaction value. Therefore, in present case, my estimates of evasion are likely to be less than actual evasion.

MVA threshold is an optimal choice for revenue administration to suppress evasion through large informal sectors. Deviation from economic neutrality is small compared to higher benefit through increased revenues. The firms do not gain any advantage for income tax purposes because they have already paid full and final liability at import stage. The response is, therefore, wholly attributable to VAT evasion. This chapter also shows that there is an increased need to study response to import stage VAT in developing countries casts doubt on the idea that VAT would self regulate evasion in low income countries.

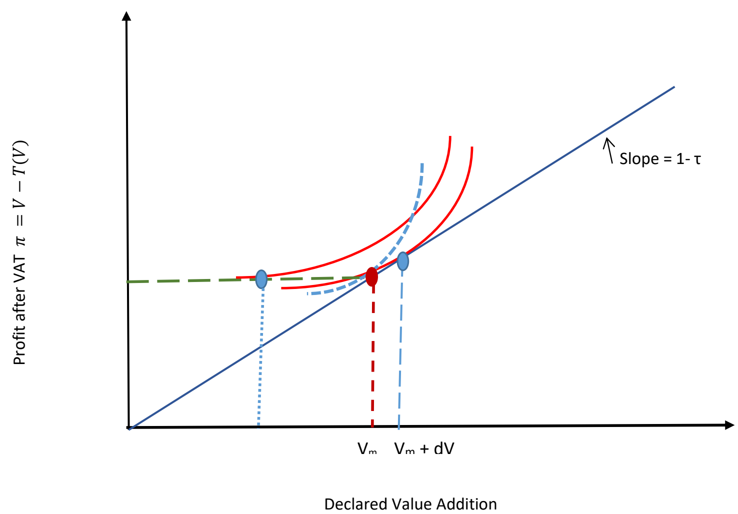
Figure 2.1: Changes in Minimum Value Addition Threshold



Explanation: This figure plots changes in Minimum Value Addition threshold for importers. In 2012, threshold jumped from 12.5% to 18.75%, a 50% increase, and also changed slightly from 18.75% to 17.65% in 2014. These changes in threshold provide the basis for studying bunching behavior of the firms.

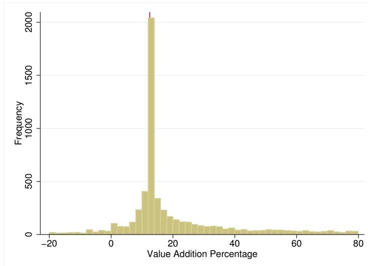
Figure 2.2: Bunching Incentive at MVA Threshold Kink

(a) Graphical Depiction of Incentive created by MVA Kink

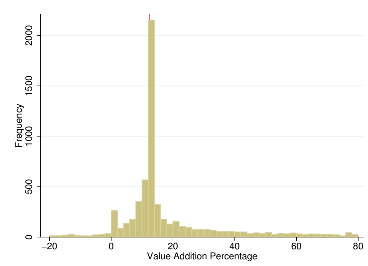


Explanation: 2(a) The figure shows bunching incentive at MVA kink. For every firm the tax rate before and after is same when there is no MVA requirement. But MVA makes pre-kink tax liability a flat tax (given by green dashed line) which the firms have to pay whether or not there actual tax liability is higher. This creates an incentive to bunch at or below this kink because this would reduce the tax liability thus increasing profit net-of-tax. The firms which have value additions in excess of MVA, shown above by $V_m + dV$, would have an incentive to bunch at or below the threshold V_m.

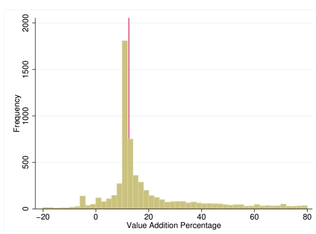
Figure 2.3: Histograms showing yearly bunching at MVA Thresholds



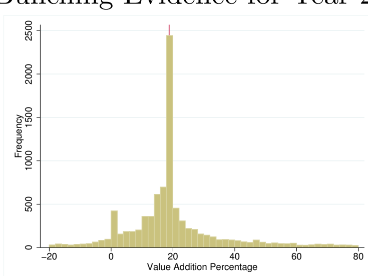
a. Bunching Evidence for Year 2009



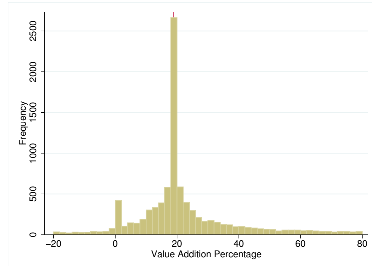
b. Bunching Evidence for Year 2010



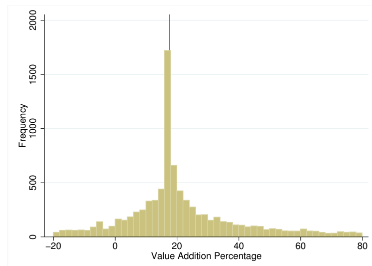
c. Bunching Evidence for Year 2011



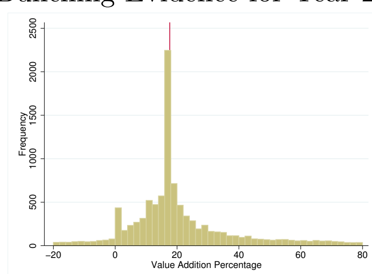
d. Bunching Evidence for Year 2012



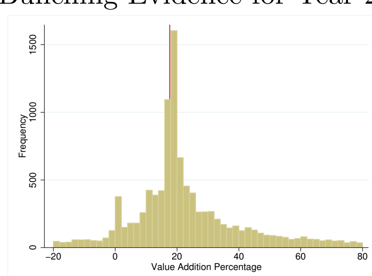
e. Bunching Evidence for Year 2013



f. Bunching Evidence for Year 2014



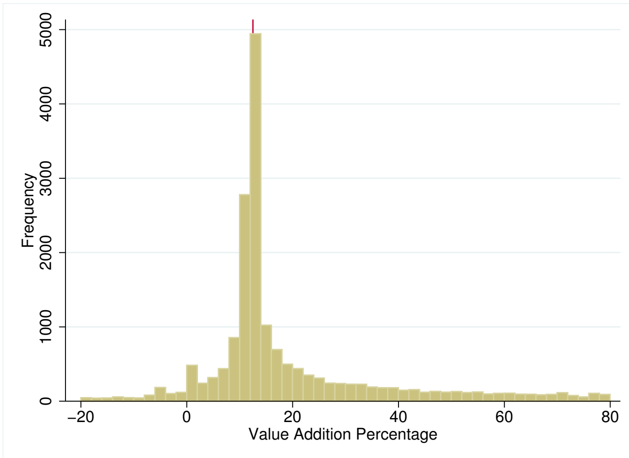
g. Bunching Evidence for Year 2015



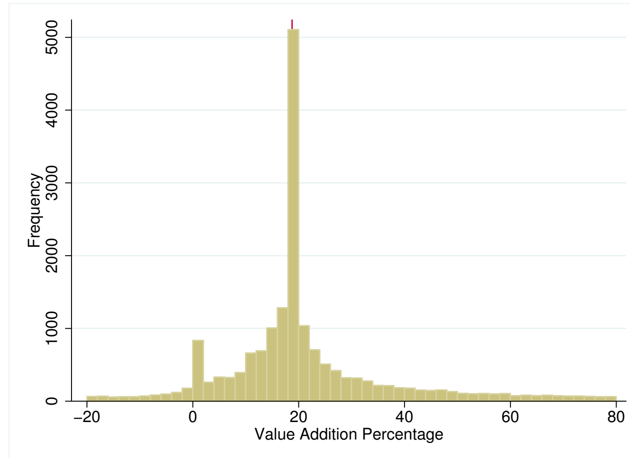
h. Bunching Evidence for Year 2016

Explanation: This figure shows bunching evidence at or around MVA threshold over eight years. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 2 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 6 percent would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. In 2012 threshold was increased from 12.5% to 18.75% (representing a 50% increase). The firms were consistently bunching around 12.5% threshold (shown by vertical red line) for years 2009-2011 but when threshold increases to 18.75% in 2012, they immediately start bunching around new threshold. In 2014, there was a very small shift from 18.75% to 17.65% but firms respond to this minimal change very aggressively by shifting to the lower bin corresponding to new threshold. It also suggests that adjustment costs are very low and firms do respond to very small benefits.

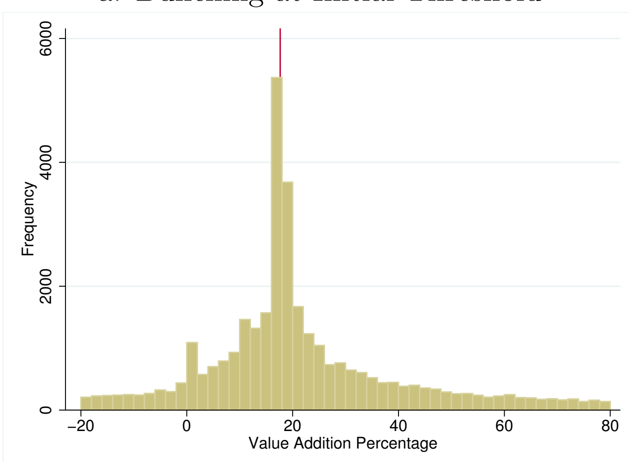
Figure 2.4: Histograms showing bunching at each MVA Threshold



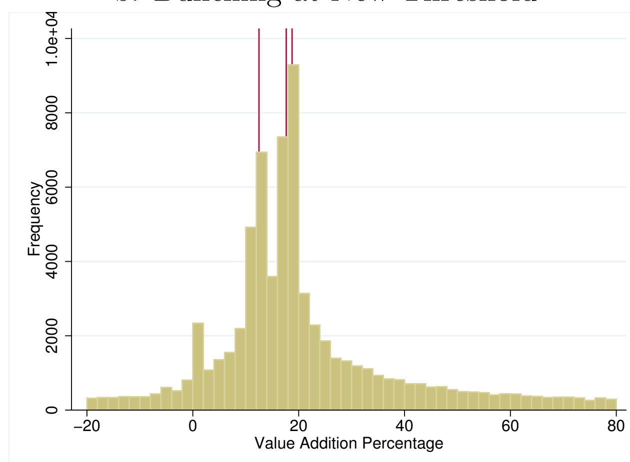
a. Bunching at Initial Threshold



b. Bunching at New Threshold



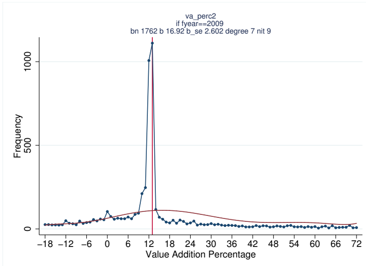
c. Bunching at Slightly Shifted Threshold



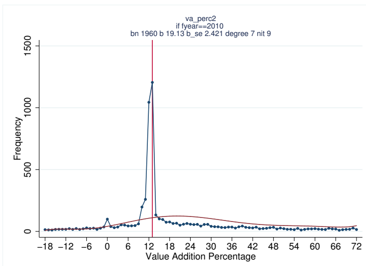
d. Cumulative Bunching at Different Thresholds

Explanation: This figure plots histograms showing bunching at different thresholds. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 2 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 6 percent would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year.Panel (a) shows cumulative bunching histogram for the initial threshold of 12.5% for years 2009-2011. Panel (b) shows a cumulative bunching histograms for years 2012-2013 when bunching threshold changed to 18.75%. Similarly, Panel (c) plots a combined histogram for years 2014-2016 when threshold changed slightly downwards to 17.65%. Panel (d) shows firm responses for all eight years with vertical red lines showing three different thresholds over time.

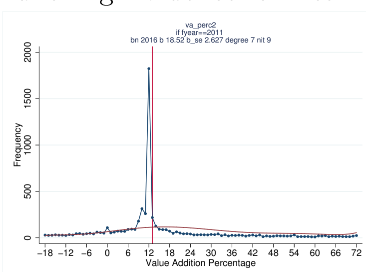
Figure 2.5: Bunching Estimates for yearly response at MVA Thresholds



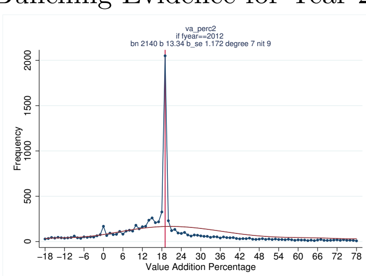
a. Bunching Evidence for Year 2009



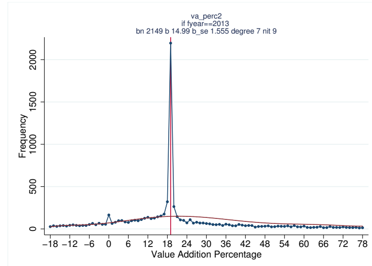
b. Bunching Evidence for Year 2010



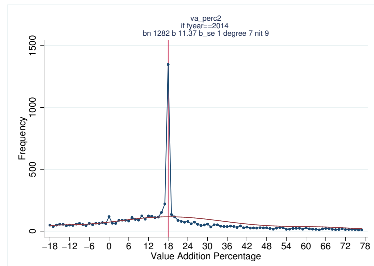
c. Bunching Evidence for Year 2011



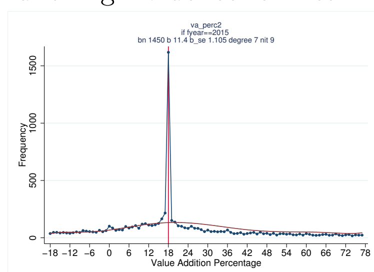
d. Bunching Evidence for Year 2012



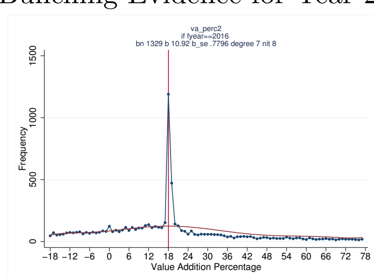
e. Bunching Evidence for Year 2013



f. Bunching Evidence for Year 2014



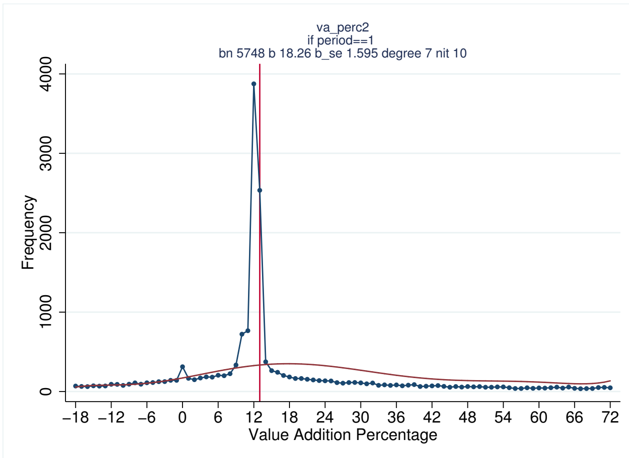
g. Bunching Evidence for Year 2015



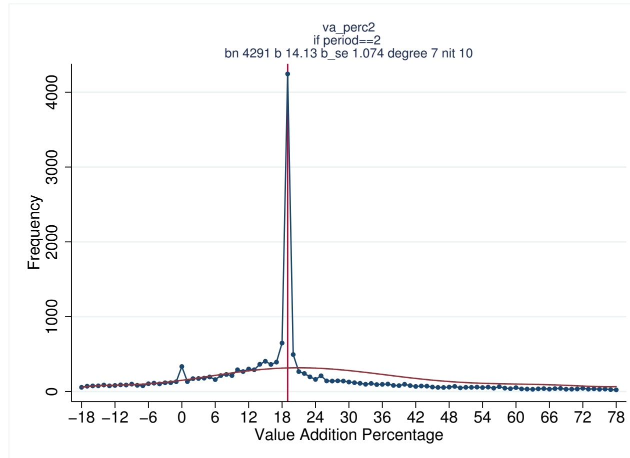
h. Bunching Evidence for Year 2016

Explanation: This figure shows bunching evidence at or around MVA threshold over eight years. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 1 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 5% percent would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. In 2012 threshold was increased from 12.5% to 18.75% (representing a 50% increase). The firms were consistently bunching around 12.5% threshold (shown by vertical red line) for years 2009-2011 but when threshold increases to 18.75% in 2012, they immediately start bunching around new threshold. In 2014, there was a very small shift from 18.75% to 17.65% but firms respond to this minimal change very aggressively by shifting to the lower bin corresponding to new threshold. It also suggests that adjustment costs are very low and firms do respond to very small benefits. The results for this are shown in Table . The bunching mass within a period when threshold is similar across years. Figure 2.6 shows bunching estimates period wise.

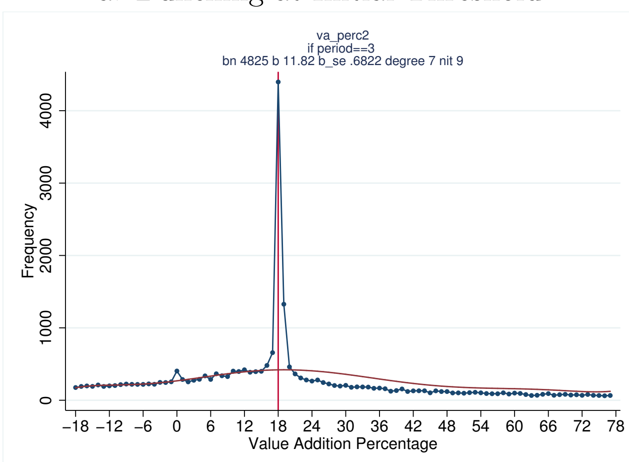
Figure 2.6: Period wise Bunching Estimates



a. Bunching at Initial Threshold



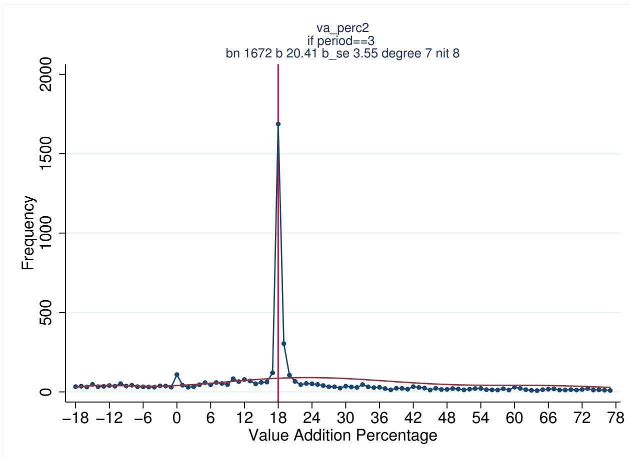
b. Bunching at New Threshold



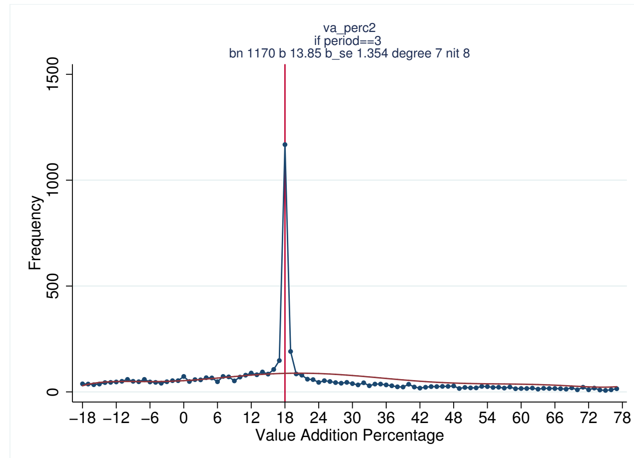
c. Bunching at Slightly Shifted Threshold

Explanation: This figure plots number of firms along with its counterfactual distribution, showing bunching at different thresholds. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 1 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 5% would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. Panel (a) shows cumulative bunching for the initial threshold of 12.5% for years 2009-2011. Panel (b) shows a cumulative bunching histograms for years 2012-2013 when bunching threshold changed to 18.75%. Similarly, Panel (c) plots a combined histogram for years 2014-2016 when threshold changed slightly downwards to 17.65%. Panel (d) shows firm responses for all eight years with vertical red lines showing three different thresholds over time.

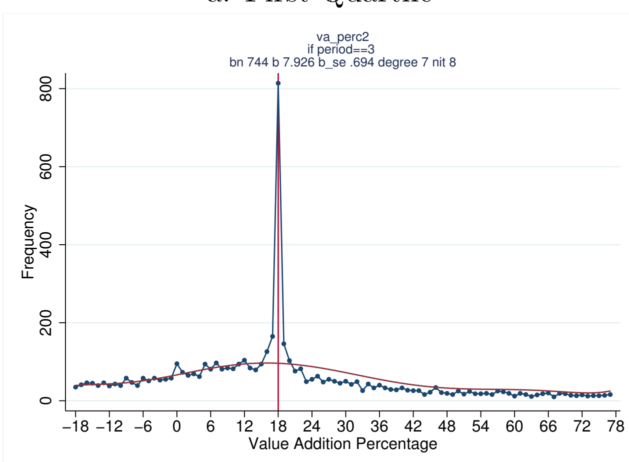
Figure 2.7: Quartile Wise wise Bunching Estimates for Third Period



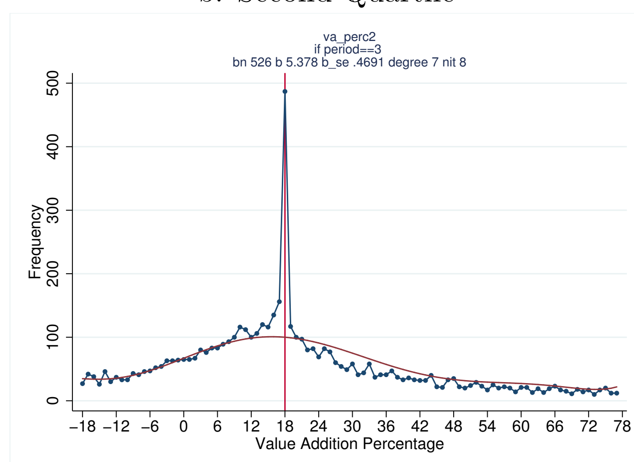
a. First Quartile



b. Second Quartile



c. Third Quartile



d. Fourth Quartile

Explanation: This figure plots number of firms along with its counterfactual distribution, showing bunching for different quartiles during third period. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 1 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 5% would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. The tails on left and right of bunching window are relatively thick for fourth quartile suggesting that these firms possibly face more adjustment costs. The results are presented at Table 2.3. For response in first and second period, see figures B.1.1 and B.1.2.

Table 2.1: Bunching Response by Year

Year	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Total Firms	5355	5727	5987	8230	7986	6306	7171	7153
Firms near Bunch Point	3323	3497	3648	4546	4298	2973	3358	3153
Percentage of firms bunching	62.05	61.06	60.93	55.24	53.82	47.15	46.83	44.08
Excess Bunching Mass, b	16.92***	19.13***	18.52***	13.34***	14.99***	11.37***	11.4***	10.92***
	(2.60)	(2.42)	(2.63)	(1.17)	(1.56)	(1.00)	(1.11)	(0.78)
Firms Bunching at Kink	2041	2189	1964	2099	2249	1373	1606	1081
Concentration	1.16	1.12	0.97	0.98	1.05	1.07	1.11	0.81

Notes: Table displays the results for bunching response by each year. Total firms is the aggregate number of firms plotted bin-wise. Firms near bunch point is number of firms within 6 bins on left side of MVA threshold. b represents the excess mass over and above the counterfactual distributions shown in figure 2.5. All values for b are statistically significant at p-value < 0.01. Bunching mass and percentage of firms bunching within 6% of bunching decline when threshold increases in 2012.

Table 2.2: Bunching Response by Threshold Period

Threshold Period	**First**	**Second**	**Third**
Total Firms	17069	16216	24389
Firms near Bunch Point	10468	8844	10944
Percentage of firms bunching	61.33	54.54	44.87
Excess Bunching Mass, b	18.26***	14.13***	11.82***
	(1.60)	(1.07)	(0.68)
Firms Bunching at Kink	6197	4348	4282
Concentration	1.08	1.01	0.89

Notes: Table displays the results for bunching response by each year. Total firms is the aggregate number of firms plotted bin-wise. Period estimate is calculated by aggregating firms falling in a certain bin by year over the period when threshold is same. Firms near bunch point is number of firms within 6 bins on left side of MVA threshold. b represents the excess mass over and above the counterfactual distributions shown in figure 2.6. All values for b are statistically significant at p-value < 0.01. Bunching mass and percentage of firms bunching within 6% of bunching decline when threshold increases in 2012.

Table 2.3: Bunching Response by Quartile for 3rd Period

Period 3	**Q-1**	**Q-2**	**Q-3**	**Q-4**
Total Firms	5172	5159	5149	5150
Firms near Bunch Point	2901	2437	2153	1993
Percentage of firms bunching	56.09	47.24	41.81	38.7
Excess Bunching Mass, b	20.41***	13.85***	7.926***	5.378***
	(3.55)	(1.35)	(0.69)	(0.47)

Notes: Table displays the results for bunching response by each year. Total firms is the aggregate number of firms plotted bin-wise. Period estimate is calculated by aggregating firms falling in a certain bin by year over the period when threshold is same. Quartiles are constructed on yearly basis, meaning that firms in one quartile for year 2014 could be in some other quartile depending on their sales volume for that particular year. Firms near bunch point is number of firms within 6 bins on left side of MVA threshold. b represents the excess mass over and above the counterfactual distributions shown in figure 2.7. All values for b are statistically significant at p-value < 0.01. Bunching mass and percentage of firms bunching within 6% of bunching decline when threshold increases in 2012.

Chapter 3

The Deterrence Value of Tax Audit: Estimates from a Randomized Audit Program

3.1 Introduction

[1]Modern tax system are based on the principle of self-assessment. Taxpayers assess their tax liability without interference from the revenue authority and report it through the tax return. The returns are considered final unless they are selected for audit. Typically, audit is the only point of contact between a taxpayer and the revenue authority and therefore the sole instrument through which the authority can punish noncompliance and create deterrence. How effectively audit does this is critical to how much revenue a country collects. (Sarin and Summers 2019) estimate that in the US around $1 trillion of additional revenue can be generated by improving IRS's audit capacity. Notwithstanding its importance to tax collection, audit has received little attention from public finance researchers. Importantly, we still do not understand fully how effective audits are in uncovering tax evasion and preventing it in future.

The central difficulty in identifying audit's role in the tax evasion decision of a taxpayer is its endogeneity. Modern tax administrations use sophisticated, risk-based algorithms to target audits toward more egregious tax evaders. While such targeting helps the authority deploy its scarce audit resources optimally, it prevents researchers from estimating audit impacts cleanly. In this paper, we overcome this central identification challenge by exploiting a national program of randomized audits from Pakistan. The program covers the entire population of tax filers in the country, and we have access to three waves of such randomized audits, leveraging which we estimate tax evasion at the baseline and audit's role in preventing it in future.

The randomized audit program began in 2013. Before that Pakistan's revenue authority (FBR) used to pick cases for audit using parametric, risk-based criteria.

[1]This chapter is prepared in collaboration with Michael Best, Email Addresses: Michael Best mcb2270@columbia.edu; and Mazhar Waseem mazhar.waseem@manchester.ac.uk.

This practice, however, was challenged before the superior courts of the country inter alia on the grounds that the criteria were confidential and likely discriminatory against some taxpayers. While these challenges were pending, the FBR could not use parametric selection and was constrained to pick audit cases using random computer ballots. It is important to emphasize that randomized audits in our setting are not a subset of audits but for three consecutive years the entire audit program of the country was randomized. We focus on VAT audits conducted under the program. The VAT return is filed every month. The high-frequency VAT data allow us to identify both immediate and distant impacts of audit on behavior cleanly.

In the standard tax compliance model, a taxpayer reports it tax liability to the government trading off the benefit and cost of tax evasion (Allingham and Sandmo 1972). The cost of evasion here is that with some probability the government would discover evasion and would recover the evaded amount along with a penalty. The probability this event occurs with is a composite term comprising the probability of audit and the probability of detection conditional on audit. In general, these two probabilities are unknown to taxpayers, although they may have formed beliefs on these based on their past interactions with the government. In our setting, the first of these probabilities is public knowledge. Before each random ballot, the FBR informed taxpayers the fraction of population to be picked for audit. The program thus creates a clean experiment whereby only the latter component of detection probability is manipulated: a random sample of firms are exposed to audit; they learn its ability to uncover evasion and update their priors accordingly. Based on the direction of such updating, they may start paying less or more revenue.

Random audits are commonly used to estimate the extent and anatomy of tax evasion in the economy. Our aim in this paper extends beyond that. We are also interested to see if audit changes the perceived likelihood of detection, thereby causing a permanent change in behavior. We do so using a long panel of administrative tax records spanning 120 months (July 2008 – June 2018), comprising the entire population of tax filers and covering both audit findings and tax returns.

We first document the results of audit. Of the 3,482 firms audited in the first wave, a positive unpaid amount was found against 986 (28.3%). In terms of volume, the unpaid amount roughly equals 8% of the aggregate baseline tax liability of *all* audited firms. For a developing country like Pakistan the evasion rate of 8% does not seem too high but its distribution is extremely unequal. The evasion rate is only around 6% for large firms (top 25%) but more than 100% for the rest. A related finding is that the former group contributes more than 99% of the revenue remitted by audited firms at the baseline. In combination, we therefore find an extreme right-skewed distribution of tax payment and a bimodal distribution of tax evasion. There roughly are two types of firms: evaders who contribute little to revenue and nonevaders who evade little and contribute roughly the entire revenue collected in the country. We obtain similar results from later audit waves.

We next look at the effects of audit on firm behavior. We have access to multiple waves of randomized audits, and our rich dataset lets us examine both proximate and distant impacts on a variety of firm outcomes. None of these impacts, however, is significantly different from zero. We examine ten intensive margin outcomes, including reported sales, costs, and revenue and one extensive margin outcome but find no effect for any of the audit waves and at any post-audit tenure. Audit seems to have no effect on firm behavior. Nor is there any heterogeneity in this result. We use two non-parametric approaches to explore heterogeneity: (1) the standard approach of adding the treatment and firm characteristic interactions into the model, and (2) the more flexible, machine-learning based approach developed in Athey, Tibshirani, and Wager (2019) using Generalized Random Forests. We divide firms on the basis of more than ten characteristics measured at the baseline including size, age, industry, location, and position in the supply chain, but find null effect in almost every subgroup we look at. Nor do we find any variation in results if we divide the sample on the basis of audit outcomes, comparing firms audit found positive liabilities against with the others or firms audited earlier with those audited later.

Pakistan's revenue authority could not audit all firms picked through random ballots. In addition, a few firms were audited by local tax offices on their own. To account for these violations of the experimental protocol, we also estimate the LATE parameters using initial random assignment as instrument. When the treatment effect is heterogeneous and there is selection into treatment on the unobserved gain, the LATE is informative only about the average effect on compliers (Imbens and Angrist 1994). To show our estimates apply to a much wider population, we use the marginal treat effects (MTEs) framework (James J. Heckman and E. Vytlacil 2005; James J. Heckman and Edward J. Vytlacil 2007), identifying a linear version of the model (Brinch, Mogstad, and Wiswall 2017; Kowalski 2016). The MTE functions we estimate are flat, showing that treatment heterogeneity and selection on unobserved gains are not important in our setting so that our LATE estimates have global external validity.

That audit produces no behavioral response means it does not reveal any new information to firms. Audit is a rare event. Only around 5% of firms in Pakistan undergo audit in a given year, meaning a typical firm experiences it once every twenty years. It is therefore surprising that audit does not register any change in firm priors in either directions. Reading this result together with the baseline distribution of tax evasion we uncover, we propose a simple explanation. Given the peculiar nature of VAT, the cost of hiding a transaction varies a lot depending on who the other party to the transaction is. If the other party is (1) a consumer, or (2) an unregistered firm, or (3) a firm willing to collude, the cost is typically low as such transactions do not produce third-party information. The cost of hiding a transaction, on the other hand, is typically high if the other party is an uncooperative firm. This results in an S-shaped detection probability function first suggested by Henrik J. Kleven et al. (2011) and later confirmed in other setting including the Pakistan's (Waseem 2019). In this world, the easy-to-detect component of the tax base is reported and the hard-to-detect component is not. Audit would change firm priors only if it goes after the latter component. Our personal interviews with auditors suggest it is usually not the case. During an audit, auditors go through returns filed by a firm line by line, verifying if each line adheres to the tax code. They, for example, see that the correct tax rate has been applied, no inadmissible input tax has been claimed, no unlawful exemption has been availed, and the tax liability has been correctly calculated. While these activities are important and are likely to result in additional revenue, they are unlikely to move firm priors on the detection probability outward.

In the existing literature, no consensus exists on the sign or magnitude of the deterrence value of audit. Earlier contributions to this line of literature are lab studies some of which do find a positive effect (see Kirchler (2007) for a survey). But in others tax evasion increases after audit (for example (Maciejovsky, Kirchler, and Schwarzenberger 2007)). This occurs either because audit forces a downward revision of the perceived detection probability or because taxpayers irrationally believe current audit makes them less likely to face future audit, a phenomenon known as the gambler's fallacy (Gilovich 1983) or the bomb crater effect (Mittone 2006). Another strand of this literature manipulates one or both components of the detection probability, sending deterrence messages to a random sample of taxpayers. To maximize power, these studies usually target more noncompliant sections of the population and their results are thus not directly comparable to ours. In a recent meta analysis covering 45 such studies, done largely in rich economies, Antinyan and Asatryan (2019) find that on average the effects of such interventions are modest, increasing the probability of compliance by only 1.5-2.5 percentage points.

Another set of studies exploit random audits to estimate their effects on future behavior. Examples include Gemmell and Ratto (2012), DeBacker et al. (2015), DeBacker et al. (2018), and Advani, Elming, and Shaw (2019). Of these, the latter two, based in the US and the UK, find significant dynamic effects of audit: the audited taxpayers continue to pay more in years after the audit. In contrast, the former two, looking at the UK taxpayers and US corporations, report a null effect. Random audits are in general not an optimal way to allocate resources by the tax authority and these audits therefore are usually a small subset of audits done in a year. This is not the case in our setting. Our sample frame is the universe of VAT filers and our randomized sample includes all audits done in a year. Our results therefore apply to a typical firm in the VAT net with the audit done under conditions (managerial oversight, intensity of audit, political economy, etc.) a typical audit would be done under. The scale of the intervention also means our estimates are robust to external validity concerns randomized studies face commonly, arising for example from ignoring the general equilibrium effects (Muralidharan and Niehaus 2017; Deaton and Cartwright 2018).

Tax evasion has received renewed research interest in recent years. This revival is driven by the strong link between the economic development and fiscal capacity of a state (Besley and Persson 2013). In part, it is also driven by the economist-as-plumber approach emphasized recently by Duflo (2017), which requires researchers to be mindful of how economic policies work in the real world. One important contribution of the paper is to use randomized audits to uncover the contours of tax evasion in a representative emerging economy. In this effort, the paper is similar to Henrik J. Kleven et al. (2011), Waseem (2019), and Waseem (2020b) who do so in other contexts. We find substantial evasion with an extremely skewed distribution. This reinforces the point in Best et al. (2015) that both economic theory and public policy must take into account enforcement constraints developing countries face more seriously than is the case now.

3.2 Conceptual Framework

This paper has two broad aims. We first use the results of randomized audits to document the extent and distribution of tax evasion at the baseline. We then examine how audit in period t affects the behavior of audited firms, in particular their tax evasion choices, in future periods. In this section, we outline a simple model that links audit to behavior, highlighting the channel through it may deter future non-compliance. The framework is based on a version of the canonical tax compliance model (Allingham and Sandmo 1972) presented in Henrik J. Kleven et al. (2011).

3.2.1 Firm Behavior to Taxation

Consider a firm that uses taxable inputs valuing $c(s)$ and nontaxable inputs valuing $\psi(s)$ to produce an amount s of output. The firm is subject to the standard VAT whereby it charges tax at the rate τ of its sales and adjusts tax paid on inputs, facing a tax liability of $T(\tau) = \tau\ (s - c)$. We assume that the enforcement is imperfect so that the firm can underreport sales $\hat{s} < s$ and overreport input costs $\hat{c} > c$, evading an amount e of its tax liability $e = \hat{T} - T$, where $\hat{T} = \tau\ (\hat{s} - \hat{c})$.

The government runs an audit program to detect tax evasion, imposing a proportional penalty at the rate θ of the evaded tax liability. The probability the government detects evasion with is $p(e)$ with $p'(e) > 0$ and $p''(e) > 0$. The firm does not know this *true* detection probability and its belief on the probability is denoted by $\tilde{p}(e)$. Based on this belief and other parameters of the tax system, the risk-neutral firm decides how much tax to evade solving the following program

$$\max_{e}\ \tilde{p}(e).\pi^{A} + (1 - \tilde{p}(e)).\pi^{NA}. \tag{3.1}$$

Here $\pi^{A} = s - c(s) - \psi(s) - \theta\tau e$ and $\pi^{NA} = s - c(s) - \psi(s) + \tau e$ denote the after-tax profits of the firm in the detected and undetected states. The FOC of the problem

$$\left[\tilde{p}(e) + e.\tilde{p}'(e)\right](1 + \theta) = 1 \tag{3.2}$$

implicitly defines the mapping between the perceived detection probability and evasion choice $e\left(\tilde{p}, \theta\right)$. The comparative statics of the problem with respect to $\tilde{p}(e)$ are unambiguous: the evaded amount decreases as the perceived detection probability increases $e(\tilde{p}', \theta) < e(\tilde{p}, \theta)$ for $\forall \tilde{p}' > \tilde{p}$.[2]

3.2.2 Audit and Belief Updating

Audit is a rare event. We show later that in a typical year the government audits only around five percent of the population, a rate at which a typical firm will experience audit once every twenty years. Audit thus represents a rare opportunity for the firm to learn the efficacy of government's detection technology, update its beliefs on it, and tailor its future behavior in accordance with the revised incentives. To see how this process works, assume that the firm's prior belief on the detection probability is a draw from the normal distribution with mean $\tilde{p}_t$ and variance $\sigma^2_{\tilde{p}_t}$. The firm undergoes

[2] See, for example, (Henrik J. Kleven et al. 2011).

audit at time t, receiving a noisy signal x_t of the real detection probability

$$x_t = p_t + \epsilon_t. \tag{3.3}$$

For simplicity, we assume that ϵ_t is also a normal process with $\epsilon_t \sim \mathcal{N}(0, \sigma^2_{\epsilon_t})$. When both the prior and signal are Gaussian, the posterior belief is also Gaussian with mean

$$\tilde{p}_{t+1} = \alpha.x_t + (1-\alpha)\tilde{p}_t, \tag{3.4}$$

and standard deviation $\sigma^2_{\tilde{p}_{t+1}} = \frac{\sigma^2_{\tilde{p}}\sigma^2_{\epsilon}}{\sigma^2_{\epsilon}+\sigma^2_{\tilde{p}}}$. The mean posterior belief is a weighted average of the signal and mean of the prior, with weights provided by the precision of each distribution

$$\alpha \equiv \frac{\frac{1}{\sigma^2_{\epsilon}}}{\frac{1}{\sigma^2_{\epsilon}} + \frac{1}{\sigma^2_{\tilde{p}}}}. \tag{3.5}$$

Intuitively, the weight $\alpha \in [0,1]$ depends on the noise to signal ratio with a more precise signal receiving a higher weight. In the extreme case, when the precision of the signal approaches infinity ($\sigma^2_{\epsilon} \to 0$), its weight tends to one and prior beliefs play no role in the formation of posterior. This simple learning model provides intuitive formulation to two conditions under which audit leads to a significant revision of firm priors on the detection probability.

Condition 1. The distribution of prior beliefs is not degenerate $\sigma^2_{\tilde{p}_t} \neq 0$.

Condition 2. The signal contains some useful information $\sigma^2_{\epsilon_t} < \infty$.

The first of these condition requires that the firm does not know beforehand the detection probability with certainty. As long as there is some randomness to the audit process, this condition must be satisfied trivially. The second condition requires that the firm gleans some new, credible information from audit. Given that audit is such a rare and intrusive process (see details in the following section), this condition must also hold. To the extent that these conditions are satisfied, they lead to the following result.

Result. If conditions 1 and 2 hold, audit causes a revision in firm beliefs on the detection probability $\tilde{p}_{t+1} \neq \tilde{p}_t$.

The revision of beliefs will in turn reflect in the firm's future behavior via the mapping $e(\tilde{p}, \theta)$. For example, in case of upward revision $\tilde{p}_{t+1} > \tilde{p}_t$, tax evasion will go down $e(\tilde{p}_{t+1}, \theta) < e(\tilde{p}_t, \theta)$ and the firm will remit more tax. To quantify the direction and magnitude of these movements, we define the deterrence value of audit (DV) as the proportional change in tax evasion caused by a marginal audit

$$DV = \frac{e(\tilde{p}_{t+1}, \theta) - e(\tilde{p}_t, \theta)}{e(\tilde{p}_t, \theta)}. \tag{3.6}$$

We call it the deterrence value because any revision of firm beliefs will impact its behavior not only in the next period but all future periods. In our empirical application, we use the variation created by the randomized audit program to estimate this deterrence value directly from the data.

Note that the functional form of the learning model we use above plays little role in our key result, although the Gaussian case simplifies the exposition considerably. Importantly, the result will hold in a general setting with $\tilde{p}_{t+1} = f(\tilde{p}_t, x_t)$ as long as the intuitive and trivial Conditions 1 and 2 are satisfied. Nor is it necessary for our result to hold that the firm must be a rational Bayesian learner. Biased learning due either to mechanical failures of inference (bounded rationality, limited attention, etc.) or to motivated thinking and cognitive tendencies of owners and managers would only mean that the updating may exceed or fall short of the rational benchmark (Bénabou and Tirole 2016). In either case, it would reflect in the firm's future behavior.

Heterogeneity. Our analysis so far is from the standpoint of a single firm. Our data, however, contain many firms which may be heterogeneous in terms of their prior beliefs as well as in how they acquire and process information and how this information maps on to their future behavior. Given that audits are randomly assigned in our sample, our empirical results capture an unbiased estimate of the *average* deterrence effect of audit. We, however, run multiple subgroup analyses to uncover any heterogeneity along these dimensions.

3.2.3 Audit Rate and Detection Probability

A common simplification in the tax compliance literature is to model the detection probability $\tilde{p}(e)$ in a reduced form way. But it is important to emphasize that this probability is a composite term comprising the audit rate (the probability that a given firm will be picked for audit) and the detection probability conditional on audit (the probability that the firm's evasion will be uncovered by audit). Denoting these two terms by $\tilde{p}_a(e)$ and $\tilde{p}_d(e)$, the detection probability $\tilde{p}(e)$ featuring in the behavioral rule (3.2) can be written as

$$\tilde{p}(e) \equiv \tilde{p}_a(e) \,.\, \tilde{p}_d(e). \tag{3.7}$$

This distinction is particularly important in our setup. Pakistan's revenue authority, before each wave of audits, explicitly announced the fraction of the population it intended to audit. With this announcement, the perceived audit rate in the population must converge toward its true value $\mathbb{E}\left[\tilde{p}_a(e)\right] \rightarrow p_a(e)$. The second component of the detection probability, however, remains unknown and only firms that undergo audit learn it from their interaction with auditors.

Of the two components of the detection probability, the existing empirical literature primarily focuses on the first. Many studies manipulate the firm's real or perceived likelihood of facing an audit through randomized interventions and examine its effects on future tax payments (see for example (Bérgolo et al. 2017) or (Slemrod 2019) for a survey). In our setup, however, all firms know the audit likelihood $p_a(e)$, but only a random subsample learn how likely the audit is to detect their tax evasion $p_d(e)$. This learning as we describe above would lead to updating of their priors, shaping the trajectory of their future tax payments.

3.3 Institutional Background

In this section, we document institutional features of the Pakistani environment that are important for our empirical analysis.

3.3.1 Randomized Audit Program

Like all tax authorities, the FBR conducts the audit of a fraction of taxpayers each year. Before 2010, the selection for audit used to take place at the local level with each regional tax office picking taxpayers from their jurisdiction for audit. In 2010, the FBR centralized this process, giving it the power to pick audits for all regional offices using a computer ballot, which could be either random or risk-based (parametric). Exercising these new powers, the FBR picked the first batch of audits using parametric criteria in 2012. The selection, however, was challenged before the superior courts mainly on the grounds that the selection criteria, which were confidential, could be discriminatory against some taxpayers. While these challenges were pending, the FBR could not pick audits using parametric criteria. The legal challenge was not resolved till the end of 2015, and during the intervening period the FBR was constrained to pick audits using random computer ballots. Importantly, random audits in our setting are not a small subsample of total audits, but for three consecutive years (2013–2015) the entire audit program of the country was randomized.

Before each random ballot, the FBR issued an audit policy that set out the proportion to be audited and the criteria for exclusion from the draw. The first information, as we note above, anchors firms' expectations on the true audit probability $\mathbb{E}\left[\tilde{p}_a(e)\right] \rightarrow p_a(e)$. The exclusions were fairly minor in the first two draws, which only excluded government departments and taxpayers already under audit. But the third draw also excluded firms under fixed and withholding type regimes of VAT. The required number of cases were picked randomly from the eligible sample (population minus exclusions) after stratifying it by business organization (corporate vs. noncorporate).[3] The ballots were held in public in the presence of taxpayer representatives, and the list of drawn cases was put on the FBR portal. The whole process was anonymous and in case was any personal information such as the name or address was revealed.[4]

The drawn cases were promptly communicated to local tax offices for initiating audits. Although these audits were conducted by the local offices, the FBR maintained central oversight through the newly developed Taxpayers' Audit Monitoring System (TAMS).[5] In addition to the centrally assigned audits, local tax offices could initiate audits on their own. But they could do so only in exceptional circumstances, such as when they received specific information on tax evasion, and only after informing the taxpayer in writing the grounds for doing so.

Table 3.1 reports descriptive statistics of the five audit waves in our sample. For our empirical analysis we use the first three only, where audit was assigned through the random ballot. The fraction of population picked (p_a) varied across audit waves, ranging between 5% and 12%. The FBR did not have the capacity to take up audits of all selected cases, and the actual audit rate in all years remained below 100% (70% for the first wave and significantly lower in the later). As we not above, local tax offices initiated a small number of audits on their own. These audits are listed in the last column of the table. Our empirical framework takes into account these two violations of the experimental protocol namely that the audit rate remained below 100% and that some audits not assigned through random ballots were conducted.

Table 3.2 shows audits were initiated soon after assignment. For example, almost 65% of those assigned through the first ballot were initiated within one month of the draw. This ratio was even higher for the later waves. Significant underpayment was

[3]Please see FBR (2015) for details of the randomization procedure, including the set of exclusions.

[4]Both audit policies and lists of drawn cases are public information and have been available on the FBR portal for view and download.

[5]TAMS was the new audit portal of the FBR. All processes related to audit, including all communications to taxpayers, were to be handled through it. This meant the FBR could monitor the progress of audits, compare it across regional offices, and take action in case of delinquency.

detected by audits. The distribution of the detected amount, however, is strongly skewed rightward, and the median detection in all three waves is zero. We present a more detailed analysis of the audit findings in section 3.5 of the paper.

3.3.2 Pakistani VAT System

Pakistani VAT largely follows the standard design. Firms charge VAT on their sales (output tax) and adjust the VAT paid on inputs (input tax). They remit the tax due (output tax minus input tax) through the tax return, which is filed every month.[6] The filing is based on the principle of self-assessment. Firms assess their own tax liability, which is considered final unless the return is picked for audit. Audit, thus, is the sole instrument through which the revenue authority can detect noncompliance and create deterrence against it.

Pakistan's revenue authority, FBR, is composed of a head office, located in Islamabad, and multiple regional office located throughout the country. These regional offices include four Large Taxpayers Units, two Corporate Regional Tax Offices and twenty Regional Tax Offices. Random audits in our sample were assigned by the head office and were completed at the regional offices. An audit team typically consists of two auditors who report to the local hierarchy. The central audit office, located at the FBR headquarter, exercises overall oversight through the online monitoring system (TAMS). Importantly, all written communications with taxpayers have to be routed through it and are considered invalid unless they contain a bar code issued by the TAMS (FBR 2015).

Revenue authorities conduct multiple types of audits, which vary in terms of their intrusiveness, such as desk audits or comprehensive audits. All random audits in our sample are comprehensive audits. In each case, the taxpayer was notified, the records were called and examined, and the results were entered into the TAMS.

Like other developing economies, tax evasion is a major issue in Pakistan. In a recent paper, Waseem (2020b) estimates an evasion rate of 35-40% among the VAT filers of the country. The tax evasion occurs through both undeclared sales and overclaimed tax credits. Given a nontrivial amount is evaded, tax audits have the potential to shift firms' beliefs on the probability of detection outward, creating deterrence against future noncompliance.

In terms of tax evasion and quality of its institutions, Pakistan is not different from other emerging economies. (Sabaini and Jiménez 2012), for example, estimate the

[6] Some small firms in some of the periods included in our sample were allowed to file on a quarterly rather than monthly frequency.

VAT evasion rate among a host of Latin American economies. These rates are quite similar to the Pakistan's.[7] Similarly, Pakistan's score on the Ease of Doing Business (59.51) is indistinguishable from the average (59.06) of all countries excluding the High Income ones (Bank 2019).[8] Nor is Pakistan an atypical country in terms of its tax morale: its score on the tax morale question in the World Value Survey is in fact better than the world average (Haerpfer et al. 2020).[9]

3.3.3 Data

We use administrative data from Pakistan that include the universe of VAT returns filed between July 2008 and June 2018. The VAT return consists of three main sections. In the first section, firms report the value of their sales, decomposing it into its foreign (exports) and domestic components. In the second section, the value of purchased inputs are reported, divided likewise in the two parts. In the final section, firms compute their tax liability, indicating the tax charged on sales, the tax credited on inputs, and the difference between the two—the tax payable. Since 2011, firms also report the transaction-level details of their sales and purchases. Each firm is assigned a unique ID and is required to file every month. The data, therefore, have a panel structure.

In addition to the return data, we use information on firm characteristics from the tax register. This information includes the business organization of the firm (corporate vs. noncorporate etc.), its date of registration, and other variables we use in our heterogeneity analysis. Appendix C.1 provides a complete list of these variables.

Finally, we use audit data available on the FBR portal and the TAMS. As we note above, the list of cases drawn in each computer ballot is publicly available. We download it from the FBR portal and merge it with our VAT return data using the unique firm ID. We are able to merge 43,465 out of 43,625 audits in our sample. For the remaining 218 cases, the firm ID mentioned in the list is incorrect. We add the audit information from the TAMS to this dataset. This information includes the date the audit was initiated, the type of audit (randomly assigned vs. locally assigned),

[7] For example, the VAT evasion rates of Guatemala, Nicaragua, Panama, and Peru are 37.5%, 38.1%, 33.8%, and 37.7%. These are withing the range for the Pakistan's estimate.

[8] The Ease of Doing Business score is widely used as a measure for the quality of institutions of a country (see for example Besley and Persson (2014)).

[9] We refer to the Question 180 on the World Value Survey 2017-2021. The question asks respondents if "Cheating on taxes if you have a chance" is justified, with responses varying from 1 (never justifiable to 10 always justifiable). Pakistan's average score on the question is 1.967, which is better than the world's average of 2.197.

and the amount detected.

3.4 Empirical Strategy

One of our empirical goals in this paper is to estimate the deterrence value of audit defined in equation (3.6). Since the VAT can be evaded by underreporting sales ($\hat{s} < s$) or overreporting input costs ($\hat{c} > c$), the DV in our setup takes the following form

$$DV = \frac{\hat{s}(\tilde{p}_{t'}, \theta) - \hat{s}(\tilde{p}_t, \theta)}{\hat{s}(\tilde{p}_t, \theta)} - \frac{\hat{c}(\tilde{p}_{t'}, \theta) - \hat{c}(\tilde{p}_t, \theta)}{\hat{c}(\tilde{p}_t, \theta)}. \tag{3.8}$$

We can compute the two terms on the RHS by estimating how reported sales and input costs respond to a tax audit, running regressions of the following type

$$y_i = \alpha + \beta\, assign_i + corporate_i + \epsilon_i, \tag{3.9}$$

where y_i is the log of reported sales or input costs, $assign_i$ denotes that firm i's audit was assigned through a random ballot, and $corporate_i$ is a dummy indicating that the firm is a corporation. For space consideration, we sometimes denote the $assign_i$ dummy simply as Z_i. Since audits in our sample are assigned randomly on stratified corporate and noncorporate samples, $\hat{\beta}$ from these regressions identifies the causal effect of interest. But most of our results are from the parallel difference-in-differences model

$$y_{it} = \mu_i + \gamma\, assign_i \times after_t + \lambda_t + \varepsilon_{it}. \tag{3.10}$$

Note that the *corporate* dummy—being time invariant—is absorbed by the firm fixed effect here.[10] This DD model offers us greater transparency (visual event-study results) and precision. We cluster standard errors at the firm level, but in some specifications we cluster at the tax office level as robustness check.

The coefficient $\hat{\gamma}$ from above model identifies the intention-to-treat effect (ITT). We also estimate the corresponding LATE parameter by instrumenting audit with initial random assignment. With treatment effect heterogeneity and selection on the unobserved gain, the LATE is informative only about the average effect on compliers (Imbens and Angrist 1994). Compliers are an interesting population in our setup. They are the firms the tax authority would audit whenever they have spare audit capacity available. Notwithstanding the policy-relevance of LATE, we are also interested to know the average effect among the population. For this reason, we es-

[10] The tax code requires a firm that changes its business organization from non-corporate to corporate and vice versa to re-register. Upon re-registration, a new identifier is issued to the firm.

timate the marginal treat effect (MTE) of audit following the framework developed in James J. Heckman and E. Vytlacil (2005) and James J. Heckman and Edward J. Vytlacil (2007). Because we have access to a binary instrument only, we cannot identify the MTE nonparametrically and do so assuming a linear functional form (Brinch, Mogstad, and Wiswall 2017; Kowalski 2016).

Table 3.3 runs balance tests on our baseline data. We compare ten VAT outcomes and ten firm characteristics across firms drawn in a given random ballot ($Z_i = 1$) with others using model (3.9). The compared groups are very similar for the first two waves: the difference in means is almost always insignificant or trivial. This, however, is not true for the third wave. Firms drawn in this wave, for example, are on average larger and more likely to be manufacturers. These differences are unlikely to have arisen by chance. We have noted in section 3.3.1 that exclusions from the draw were significantly expanded for the third wave. Importantly, firms under fixed and withholding regimes were excluded from audit. We do not identify these firms in our data and are thus unable to replicate the sample used for the random ballot of the third wave. For this reason, we focus solely on the first two waves for our empirical results. Nonetheless, for the sake of completeness we always present our main results for the third wave as well.

3.5 Tax Evasion at the Baseline

Audits we consider are randomly assigned. The amount detected by them therefore represents an unbiased estimate of tax evasion at the baseline. In this section, we document the average amount detected by audit, examining in particular its relationship with firm observables.

Table 3.4 presents the results. All amounts in this table are in PKR billions. The top row shows that 3,482 firms were audited in the first wave. These firms reported a total turnover of around 500 billion in the baseline year. The audits detected 2.15 billion of short payment against them, which constitutes 0.45% of the turnover. These firms remitted 28.16 billion of VAT at the baseline with an average effective tax rate of 5.65% (columns 5–6). The unpaid revenue therefore amounts to nearly 8% of the reported tax liability (column 7).

The next five rows of the table decompose the average rate. The second row shows that positive liability is detected against 28% of firms. The detected amount equals two-thirds of the VAT remitted by these firms. The next four rows divide firms into four quartiles based on their annual turnover in the baseline year. Strikingly, the detected amount exceeds reported tax liability for all the bottom three quartiles, implying an evasion rate of over 100%. In contrast, the evasion rate is only 6% in the top quartile. The top-quartile firms also contribute disproportionately to the tax revenue. Of the 28.16 billion VAT remitted by the audited firms, more than 99% (27.91 billion) was remitted by them. We find qualitatively similar results for the second audit wave, although the evasion at the top is even lower for this wave.

Figure 3.1 examines the relationship between tax evasion and firm size more deeply. We divide audited firms into 10 or 20 groups based on their annual turnover at the baseline and see how the evasion rate and tax payments vary with firm size. Tax evasion is particularly high at the bottom; it then declines almost monotonically before falling sharply at the top. The government revenue as a result comes almost exclusively from firms at the very top. These results are not surprising. Recent models of tax compliance in weak enforcement setting predict such a distribution of tax evasion;[11] although to our knowledge we are the first to document this stark pattern empirically. Large firms tend to have transparent accounting mechanisms within the firm. These mechanisms let them operate at their economically optimal scale, but render commonly used strategies to evade taxes—such as cash payments or keeping double books of account—infeasible.[12] Tax evasion as a result is lower among large firms which end up remitting a disproportionate chunk of revenue.

In the audit data, the detected amount is reported in six heads. Table C.2.1 decomposes the detected amount into its major heads. Less than 2% of the detected amount is recovered at the time of audit either by direct payment (column 2) or by curtailing the taxpayer's refund claim (column 7). The rest of the amount being subject to quasi-judicial adjudication and appeal processes can be recovered only after these processes have run out. We do not have data on the outcome of these processes but anecdotal evidence suggests they are cumbersome and inefficient so that the detected amount remains stuck in litigation for a long time.[13]

Although audits in our sample were randomly assigned, the audit rate for both waves remained below 100%. If audits were targeted toward specific firm types, selection resulting from it could bias the evasion rates we report above. Figure C.2.1

[11] See for example (Henrik Jacobsen Kleven, Kreiner, and Saez 2016; Gordon and Li 2009; Kopczuk and Slemrod 2006).

[12] Without strong internal controls, firms cannot grow beyond a given scale as they may worry about pilferage and stealing by local managers.

[13] According to a recent press report a total of 76,700 cases involving a recoverable amount of PKR 1.77 trillion are stuck in litigation. Nearly two-thirds of the litigated amount (PKR 1.1 trillion) is pending internally (at the two appeal fora available within the FBR) and the rest with the superior courts of the country. For details of these numbers see here.

explores such selection, examining if firms audited early were systematically different from those audited later. We find no systematic correlation between the amount detected and the order in which audits were taken up. Nor is the order correlated with other firm observables (see Table C.2.3). A much detailed analysis of selection appears later in the paper. We find no evidence of such selection: within the randomly assigned sample, audits do not appear to target any specific group. To this extent, our estimates represent unbiased estimates of noncompliance at the baseline.

Tax audits are unlikely to uncover all tax evasion. For this reason, revenue authorities that use random audits to estimate the tax gap multiply the detected amount by a scale factor to convert it into their official estimate. IRS, for example, uses a scale factor of 3.28 for this purpose. The factor is derived from a direct survey of taxpayers on tax compliance (see IRS 1996; Henrik J. Kleven et al. 2011 for details). We do not have access to such a multiplying factor for the VAT in Pakistan. Nor are audits in our sample *extensive* audits, done for the express purpose of measuring noncompliance. They rather are routine audits revenue authorities conduct during the course of their normal operation. Our estimates therefore likely represent a conservative lower bound on the true evasion rate in Pakistan.

3.6 Audit and Firm Behavior

We now examine the effects of audit on firm behavior, assessing in particular if they deter tax evasion in future periods.

3.6.1 ITT Estimates

We begin by presenting nonparametric evidence. Figure 3.1 plots the coefficients δ_js from the following regression

$$y_{it} = \mu_i + \sum_{j=2}^{N} \delta_j .\, 1.(month{=}j)_t + u_{it}, \tag{3.11}$$

where y denotes the log of variable indicated in the title of each panel. The regression is run separately for firms drawn in the random ballot ($assign_i = 1$) and other firms in the sample ($assign_i = 0$).[14] We drop the dummy for the first month (July 2008) and plot coefficients on the other month dummies (up to June 2018). Figure 3.4 illustrates

[14] The sample here includes all firms other than government departments and firms already under audit. Both categories of excluded firms together constitute a small (<5%) fraction of the $assign = 0$ sample.

the DD version of these plots, where we add interactions of the month and *assign* dummies into (3.11) and plot the coefficients on these interactions along with the 95% confidence intervals around them. Given the drawn firms are a random sample of the population, it is unsurprising that the trajectory of treated and untreated outcomes is indistinguishable from each other in the 62 pre-draw months. Table C.2.4 shows this formally by estimating baseline trends using model (3.10).

Strikingly, however, the outcomes continue to evolve on the common, preexisting trend even in the post-draw period. The relative difference between the two groups remains indistinguishable from zero in the 70 post-draw months we consider. Figures 3.5 and 3.6 replicate this analysis for the second draw, showing similar results. Initial evidence thus suggests that audit does not cause significant revision in firm priors on the detection probability and thus does not induce a significant change in behavior. Below, we examine this result in more details by running formal, regression-based tests.

The top panel of Table 3.5 reports our ITT estimates from model (3.10). We examine both short- (one-year) and medium-run (three-year) impacts produced by the audits assigned in the first wave. Consistent with the visual evidence none of the ten coefficients differs significantly from zero at the conventional level. Nor is there any systematic difference between the proximate and distant responses. Table 3.6 repeats the exercise for the second wave. Tables 3.7–3.8 examine six other VAT outcomes, and Table C.2.5 clusters at the tax office level. All these 46 specifications—covering ten intensive margin outcomes, one extensive margin outcome, and two audit waves—tell a consistent story: audit does not have a meaningful impact on firm behavior, either in the short or in the long run.

3.6.2 LATE Estimates

Since the FBR did not conduct audit of all cases drawn in the random ballots, the above estimates capture the average effect of getting *picked* for audit rather than the average effect of audit. To compute the latter parameter, we estimate the 2SLS models corresponding to (3.10), instrumenting the endogenous variable *audit* by the initial random assignment.[15] Table C.2.6 reports the first stage of these regressions, illustrating that a strong first stage exists in this setting. The bottom panels of Tables 3.5–3.8 and C.2.5 report the LATE estimates for the 46 specifications we run. The results are similar. The majority of the LATE estimates are of negative sign, statistically insignificant, and economically trivial.

[15] For brevity, we sometimes denote $audit_i$ variable simply as D_i in the subsequent sections.

Figures C.2.2-C.2.3 and Table C.2.7 examine the third wave of audits, reporting parallel results comprising the ITT and LATE estimates. Recall that for this wave the balance tests reveal significant differences between $Z_i = 1$ and $Z_i = 0$ groups (see Table 3.3). We therefore do not draw any conclusion from these results and produce them only for the sake of completeness.

3.6.3 ATE Estimates

When treatment effects are heterogeneous and there is selection into treatment on the unobserved gain, the LATE is informative on the average effect of the treatment on compliers only (Imbens and Angrist 1994; Abadie 2003). Compliers, in our setting, are firms that are pushed into audit by the instrument (being drawn in the random computer ballot). The LATE we identify therefore may not reflect the average effect in the population unless the impact of audit does not vary across firms or auditors do not target specific firms, using information we do not observe.

We first explore the latter point, examining if auditors target selective types of firms. Table 3.9 compares audited and unaudited firms.[16] Audited firms here include both that were picked by a random draw ($Z_i = 1$) and that were picked by local tax offices based on their information ($Z_i = 0$). Tables 3.10-3.11 separate the analysis for the two subgroups. A typical audited firm indeed differs from the unaudited in terms of observables we examine (Table 3.9). But these differences are almost entirely driven by the small group of firms local tax offices picked for audit on their own ($Z_i = 0$). Within the random-assignment group ($Z_i = 1$), audits do not seem to target any selected subgroup. Figures 3.7-3.8 compare audited and unaudited firms in our event study framework (3.11). Since the specification includes firm fixed effects, the results capture any residual selection into audit which is not explained by the firm's fixed characteristics, such as size or industry. There does not appear to be any such residual selection as the reporting histories of both groups are similar. Table C.2.8 establishes this rigorously by running formal tests on the baseline data. Parallel trends for a long preaudit period mean our DD estimator remains internally valid and applies to all audited firms rather than compliers only.

The above result is supported by our two previous results. First, the compliance rate falls from 70% in the first audit wave to 30% in the second, yet we see no meaningful difference between the corresponding LATE estimates (compare Tables

[16] Since audits were done at the local tax office, we need to compare audited and unaudited firms within a tax office to rule out selection. We therefore include tax office fixed effects into these regressions.

3.5 and 3.6). This suggests that the marginal firm pushed into audit may not be significantly different from others within the randomly assigned ($Z_i = 1$) sample. Second, the amount detected and other firm observables bear no correlation with the order in which audits were taken up (Figure C.2.1 and Table C.2.3). This suggests that audits are not systematically targeted toward specific group of firms. Auditors do not seem to possess any privileged information to do so.

Continuing our effort to go beyond LATE, we next exploit the marginal treatment effect (MTE) framework popularized by J. J. Heckman and E. J. Vytlacil (1999). Since our instrument is binary, we cannot identify the MTE function nonparametrically and instead identify a linear version of it following Brinch, Mogstad, and Wiswall (2017) and Kowalski (2016). Figures 3.9–3.10 show the MTEs we estimate using the two randomization waves as instruments. The technical details of the estimations are in Appendix C.2. Importantly, the MTEs from all specifications are flat. The change in the unaudited outcomes as the potential fraction audited increases reflects selection. On the other hand, the gradient in the audited outcomes reflects selection and audit effect heterogeneity. That both these curves are flat rules out these factors in our setup. Note that the functional form assumption we make is not too restrictive. We have access to two randomized experiments and therefore can exploit more information than is typically available in an RCT. Specifically, because the compliance rate varies between the two waves, both audited and unaudited outcomes in our setup are identified at four rather than two points. The flat MTEs we obtain from all specifications therefore suggest that our LATEs have global external validity.

3.6.4 Heterogeneity

To strengthen the above conclusion, we also examine treatment heterogeneity directly. We do so using two nonparametric approaches. First, we estimate triple-difference versions of model (3.10), interacting the DD term with firm traits. We explore eight traits introduced into the model as dummies indicating (i) firm size; (ii) firm age; (iii) firm location; (iv) local tax office having jurisdiction over the firm; (v) the type of local tax office (LTU vs. RTO etc.); (vi) firm's position in the supply chain (manufacturer vs. wholesaler etc.); (vii) firm's business organization; and (viii) industry the firm operates in. All these traits are measured at the baseline before the announcement of ballot results, and we estimate the model separately for the two audit waves. Figures C.2.4-C.2.11 display the results. We do not find any systematic treatment effect heterogeneity across the subgroups we compare. The 95% confidence interval almost always includes zero, showing that the response of each subgroup is statistically indistinguishable from that of the omitted category.

In addition to the predetermined firm traits, we also explore heterogeneity by the timing and outcome of audit. Figure C.2.12 divides audited firms into ten groups, depending upon the time lag between the assignment and initiation of audit. If auditors have hidden information they use to target specific subgroups, it would be reflected in the order they took up the assigned audits in. We, however, do not see any differences along this dimension. Audited firms in all deciles appear to be very similar. Table C.2.9 stratifies the audited sample by the detected amount, looking for any differential effect upon firms auditors did find an underpaid amount against. Here also we do not find any differential effect.

Finally, we explore treatment heterogeneity using a more flexible machine-learning approach. We ask if the audit effect varies with the firm's predetermined traits using the Generalized Random Forest algorithm developed in Athey, Tibshirani, and Wager (2019).[17] To reduce the computational demands of the algorithm, we use the simple difference-in-means model (3.9) as the baseline rather than the DD model (3.10) we have been using so far. The results are in Figures C.2.13-C.2.22. The first four of these figures show the audit effect does not vary with firm size or age. The rest of the figures explore binary traits. Again, we do not find any systematic heterogeneity in the audit effect along any of the eight traits we look at.

3.7 Why Audit has No Effect on Behavior?

We present extensive evidence above showing that audit does no affect future firm behavior. Not only does this finding hold on average but also among more than 20 subgroups we define based on firm observables. In terms of our model, these results suggest that either Condition 1 or Condition 2 fails. This is surprising given that these conditions are so trivial and intuitive. The failure of the first condition, for example, implies that prior beliefs of all firms are concentrated on the true detection probability with no variance around the mean. This notion is extremely unlikely as existing evidence shows that taxpayers misperceive even the most simple and accessible details of the tax system, such as the marginal tax rates (Taubinsky and Rees-Jones 2016). How can then they be expected to know something that has not even been revealed and that too with certitude. The failure of the second condition is also equally

[17] In the approach, individual trees are grown by greedy recursive partitioning of the sample space, with each split chosen to improve the model fit. The trees are then randomized using bootstrap aggregation, whereby each tree is grown on a different random subset of the training data, and random split selection that restricts the variable available at each step of the algorithm.

unlikely. Not only is audit a rare event,[18] it is quite intrusive as well. Auditors spend considerable time with taxpayers going through their records, visiting their premises, and discussing audit findings. It is therefore highly unlikely that taxpayers do not glean any useful signal on the government's detection technology during this weeks-long interaction. No updating in either directions and consequently no reoptimization of future behavior is puzzling. In this section, we make sense of this result.

We begin by tweaking the model we presented in section 3.2 slightly. Following Basri et al. (2019), the revised model treats evasion as a discrete rather than the continuous choice. Discretizing the choice variable brings the model closer to our VAT setting, leading to simpler and more intuitive exposition. The firm engages in L transactions, indexed by $l = 1...L$, and decides separately for each transaction whether to report or hide it. It would report a transaction and remit the VAT due if the cost of hiding it exceeds the benefit

$$\left[\tilde{p}_l(e_l) + e_l.\tilde{p}_l^{'}(e_l)\right](1+\theta) > 1. \tag{3.12}$$

This inequality is a discrete version of the behavioral rule (3.2), showing that the firm's choice critically hinges on the detection probability hiding a transaction entails. Ordering transactions in terms of the hiding cost, we can define L^* as the first transaction for which the above inequality holds. The firm will accordingly report transactions $L^*...L$ and will remit the tax due, amounting to $\int_{L^*}^{L} \tau(s_l - c_l)\, d(l)$. Note that L^* could be the first transaction, in which case the firm does not evade at all, or it could be the last, in which case the firm evades the entire tax due. In general, L^* would be idiosyncratic to the firm, depending on its scale, trading network, and other characteristics.

Note that hiding a transaction would be easier for the firm if the other party to it is (1) a consumer, (2) an unregistered firm, or (3) a firm willing to collude. In these cases, the firm can cover its tracks, making it harder for the government to detect evasion. On the other hand, hiding a transaction would be difficult if the other party is unwilling to collude, such as a firm that cannot handle unaccounted cash and therefore cannot keep a transaction out of books.[19] The $\tilde{p}_l(e_l)$ faced by the firm on

[18] During the ten-year period we consider, the FBR could not audit more than 5% of firms a year, a rate at which a typical firm would experience audit once every twenty years. Note that the likelihood of a firm facing the audit is endogenous to firm behavior if the authority runs a parametric, risk-based system of audit selection. The raw audit probability is for illustrative purpose only, showing that on average the authority can only audit one-twentieth of the population each year.

[19] These consideration can lead to segmentation of firms into good and bad VAT chains with compliant firms dealing with compliant firms only and vice versa. See Paula and Scheinkman (2010),

different transactions therefore takes the shape shown in Figure 3.11. It is typically low for the former type of transactions but turns sharply once the latter type begins. Such an S-shape detection probability function was first suggested by Henrik J. Kleven et al. (2011) and has since then confirmed in other empirical settings (see Waseem (2019) for one such example). The shape reflects that the probability of detection to a first order depends on the external information an economic transaction generates for the government.

The discrete choice model predicts a simple behavioral rule. The firm will report transactions entailing high detection probability $[L^*, L]$, hiding the rest. In this world, it is easy to see why audit may not cause any *observable* change in future behavior. For this purpose, let us characterize a marginal audit as the following.

Definition. An audit is *pivotal* if it leads to the flipping of inequality (3.12).

A pivotal audit causes sufficiently large revision in the firm's perceived detection probability so that the LHS of inequality (3.12) exceeds one after the audit if it did not do so earlier and vice versa. For example, indexing the post- and pre-audit variables by $t+1$ and t, an audit will be pivotal if

$$\begin{aligned} \left[\tilde{p}_{l,t}(e_{l,t}) + e_{l,t}.\tilde{p}'_{l,t}(e_{l,t})\right](1+\theta) &< 1 \\ \left[\tilde{p}_{l,t+1}(e_{l,t+1}) + e_{l,t+1}.\tilde{p}'_{l,t+1}(e_{l,t+1})\right](1+\theta) &> 1. \end{aligned} \tag{3.13}$$

In this case, the transaction l will not be reported prior to audit but will be reported after it. Thus, a necessary condition for audit to cause an *observable* change in firm behavior is that it is pivotal.

If the structure of the detection probability is of the form shown in Figure 3.11 with the probability on most transactions being close either to zero or one, it is highly unlikely that a given audit will be pivotal. Indeed, in this world even when Conditions 1 and 2 are satisfied so that audit does cause a revision in firm priors, it would still not cause an observable change in firm behavior if such revision is not large enough to flip inequality (3.12). This is particularly likely if the auditors do not have incentives or resources to uncover transactions not reported by firms $[1, L^*)$. They, for example, may have a shorter planning horizon than the government and therefore may not value the dynamic gains from uncovering hidden transactions as much as a forward-looking planner will do. Instead, they may focus on already-

Gadenne, Nandi, and Rathelot (2019), and Gerard et al. (2019) for empirical evidence on market segmentation caused by a VAT.

reported transactions only, uncovering any mechanical violations of the tax code and therefore the underpayments of tax.

In our data, we do not observe activities performed by auditors during an audit, but personal interviews with them reveal that they indeed spend most of their time checking mechanical violations of law. During an audit, they go through the returns filed by a firm line by line, verifying if each line adheres to the tax code, ensuring for example that the correct tax rate has been applied, no inadmissible input tax has been claimed, no unlawful exemption has been availed, and the exact tax liability has been calculated. While these activities are important and are likely to bring additional revenue, they are unlikely to cause sufficiently large revision in firm priors that would reflect in their future behavior as a deterrence effect of audit.

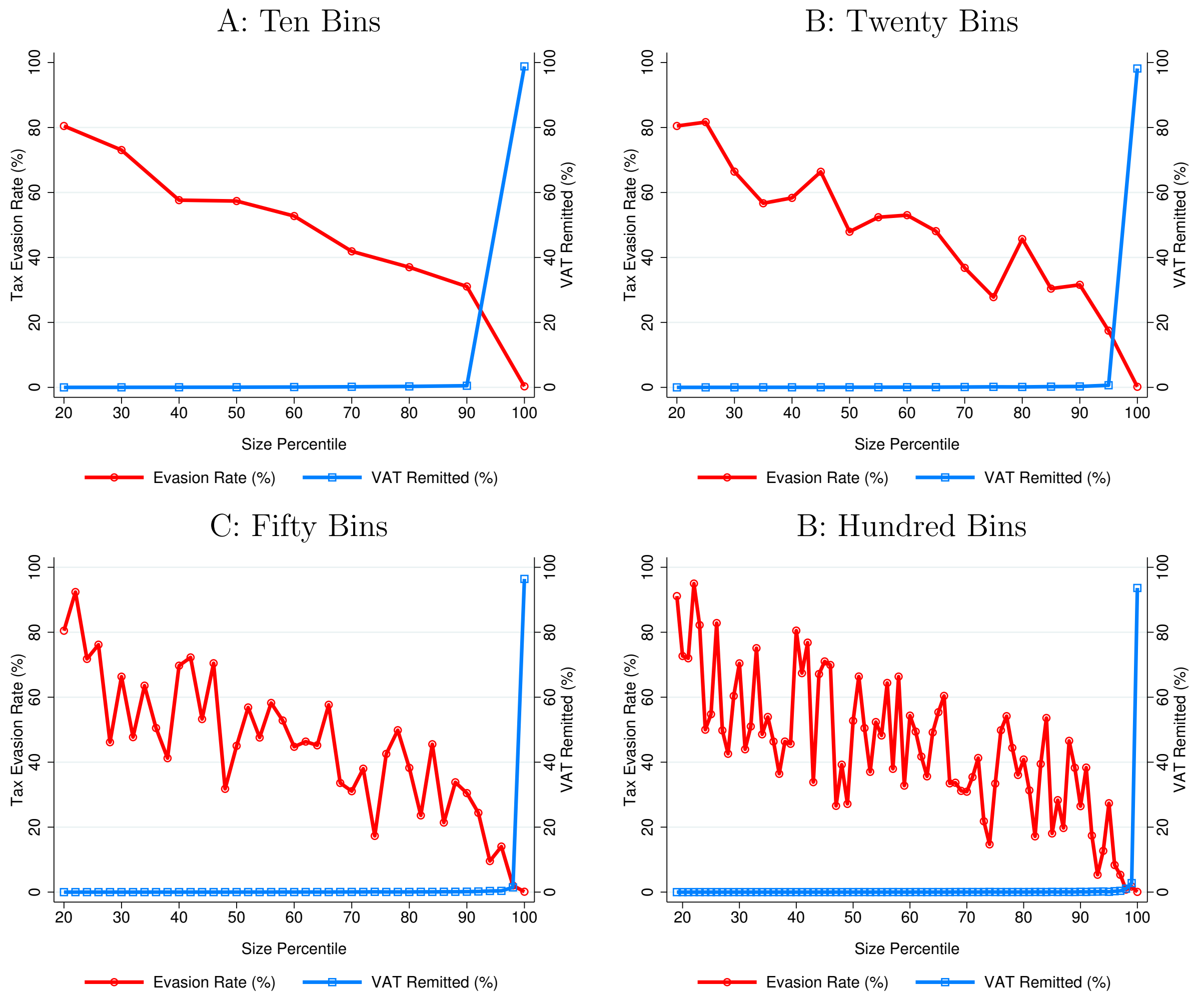
A testable prediction of the proposition that auditors devote little attention to transactions not reported by firms is that the detected amount will fall as the proportion of such transactions in a firm's sales rises. Table C.2.10 tests this prediction. We divide firms into four groups based on the share of final sales in their turnover at the baseline. Final sales are transactions where the other party is either a consumer or an informal firm, and we identify these using our transaction-level data. Theory predicts that the likelihood of not reporting transactions of such a kind is much higher than others, and audit therefore must detect a greater amount against firms a greater proportion of whose turnover comprises such transactions. But it is not what we find. The amount detected in fact falls as the share of final transactions in a firm's turnover rises. This correlation holds even when we add important covariates to the specification including firm size. The evidence thus supports the notion that much of the audit effort goes into reconciling the already-reported transactions rather than uncovering the unreported ones.

3.8 Conclusion

In modern tax systems, audit is to some extent the sole instrument through which the revenue authority can detect and deter tax evasion. We exploit a national program of randomized audits from Pakistan to examine how much evasion audit detects and how much evasion it prevents by changing post-audit behavior. Combining VAT returns and audit outcomes data, we find audit detects a substantial amount of evasion: the detected amount is 8% of the aggregate annual turnover of audited firms. The evasion rate, however, varies substantially across firms. It is more than 100% among firms in the bottom three size quartiles but only 6% among the rest. Despite detecting such a large amount of evasion, audit does not create any deterrence against it. Examining more than ten intensive and extensive margin outcomes, we find no significant impact of audit on immediate or distant behavior for any of the randomization wave we consider. This result is robust to a number of specification checks, and we do not find any heterogeneity in audit effects across any subpopulation.

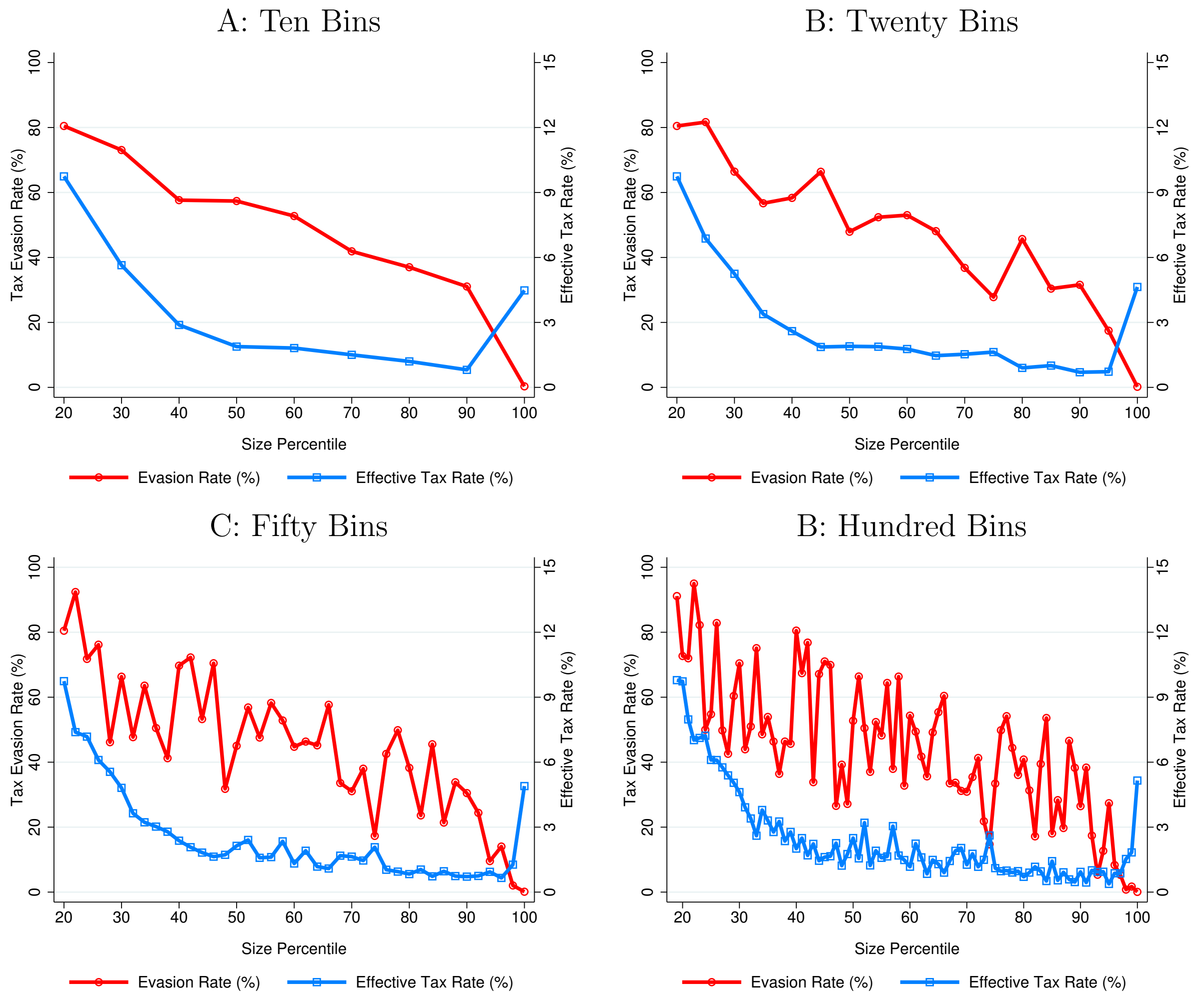
That audit does not affect behavior is puzzling. Audit is a rare event, with a typical firm likely to experience it once every twenty years. Lack of response to it means audit does not reveal any new information to firms. We suggest a simple explanation of this result. Transactions carried out by a firm can be roughly divided into two types. Transactions with consumers, unregistered firms, or colluding firms can be hidden easily, while those with uncooperative firms cannot. In this world, profit-maximizing firms report easy-to-detect transaction but hide the rest, and audit would change firm priors only if it goes after the hidden transactions. Our interviews with auditors reveal it is usually not the case. Instead, auditors scrutinize reported transactions only, looking for any mechanical violations of law. Insufficient focus on uncovering hidden transactions means audit does not change firm priors on the detection probability and thus does not induce a permanent change in behavior.

Figure 3.1: Evasion Rate By Firm Size



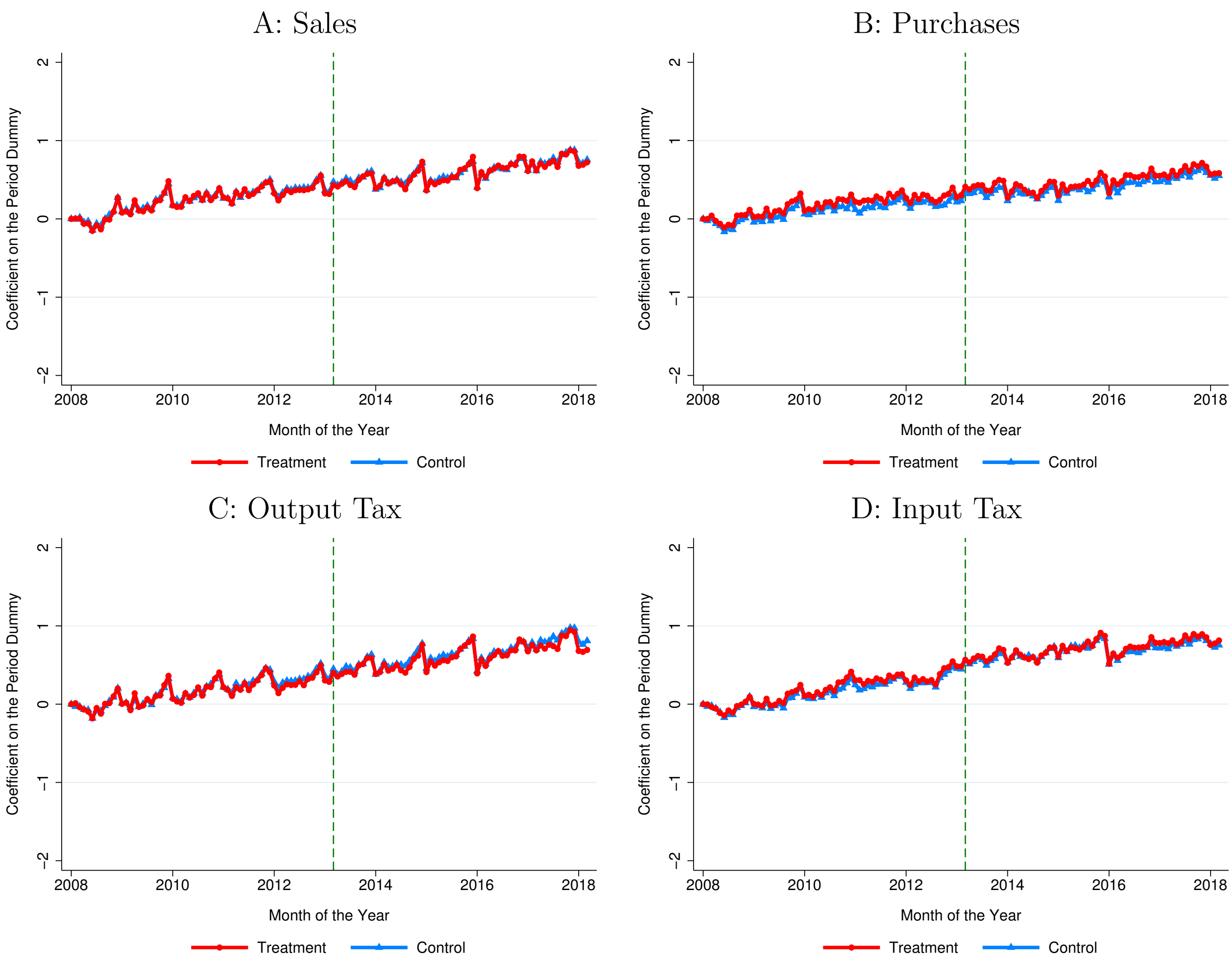
Notes: The figure plots the tax evasion rate by firm size. In the top panel, we divide firms into 10 equal groups based on their annual turnover in the baseline year. We calculate the evasion rate in each group as the total amount detected by audit against all firms in the group as a fraction of total VAT remitted by these firms at the baseline. This evasion rate is shown by the red curve with the y-axis on the left. To maximize power, the sample here includes all firms audited in the first two audit waves. We superimpose a series indicating the total VAT remitted by firms in each group as a fraction of total VAT remitted by all firms in this sample. This series is shown by the blue curve with the y-axis on the right. The bottom panel repeats the exercise after dividing firms into 20 equal groups on the basis of their baseline turnover. Both plots begin from the 20th percentile because firms below this threshold remit no VAT at the baseline so that their evasion rate is not defined.

Figure 3.2: Evasion Rate By Firm Size



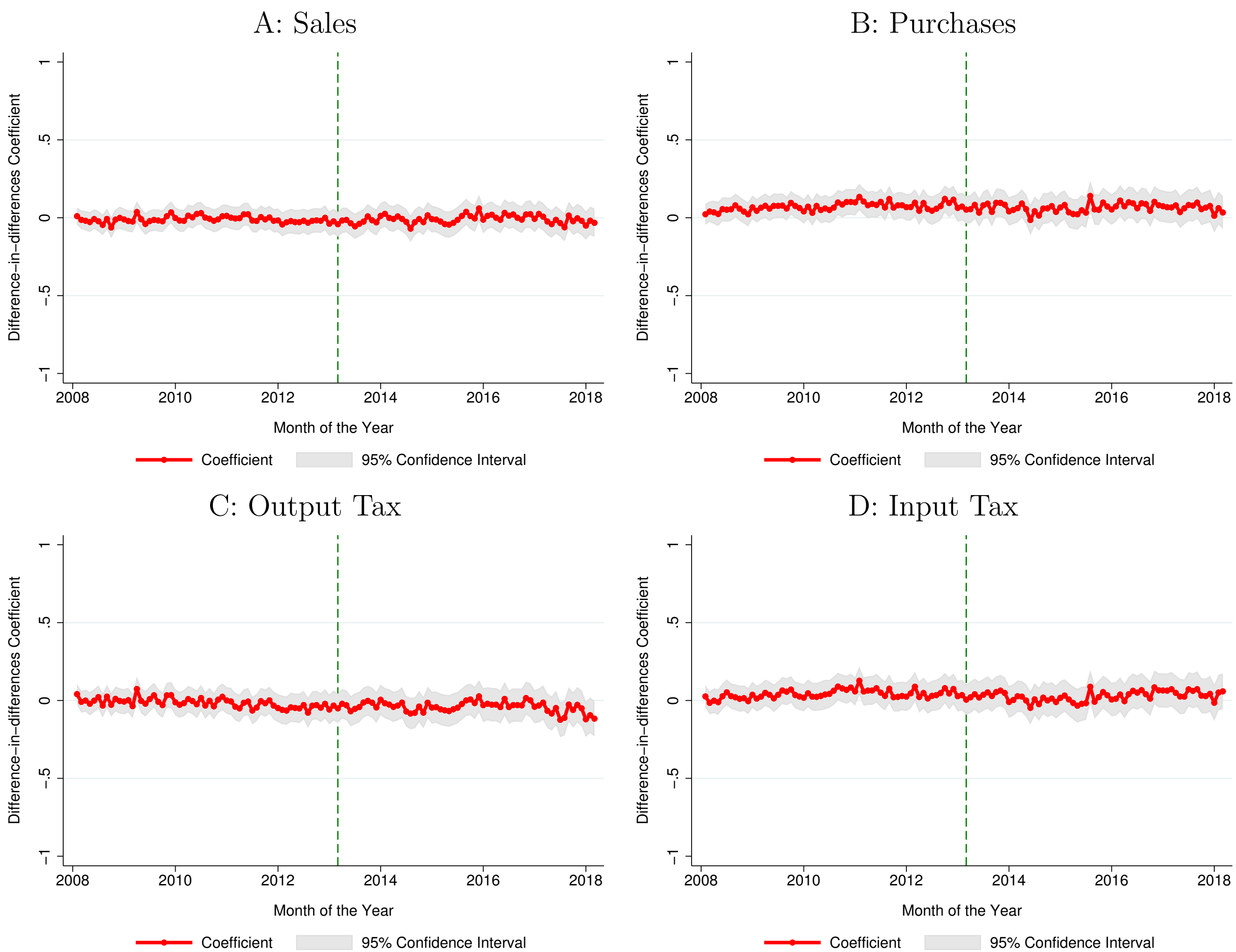
Notes: The figure plots the tax evasion rate by firm size. In the top panel, we divide firms into 10 equal groups based on their annual turnover in the baseline year. We calculate the evasion rate in each group as the total amount detected by audit against all firms in the group as a fraction of total VAT remitted by these firms at the baseline. This evasion rate is shown by the red curve with the y-axis on the left. To maximize power, the sample here includes all firms audited in the first two audit waves. We superimpose a series indicating the total VAT remitted by firms in each group as a fraction of total VAT remitted by all firms in this sample. This series is shown by the blue curve with the y-axis on the right. The bottom panel repeats the exercise after dividing firms into 20 equal groups on the basis of their baseline turnover. Both plots begin from the 20th percentile because firms below this threshold remit no VAT at the baseline so that their evasion rate is not defined.

Figure 3.3: Intention to Treat Effects of Audit – First Wave



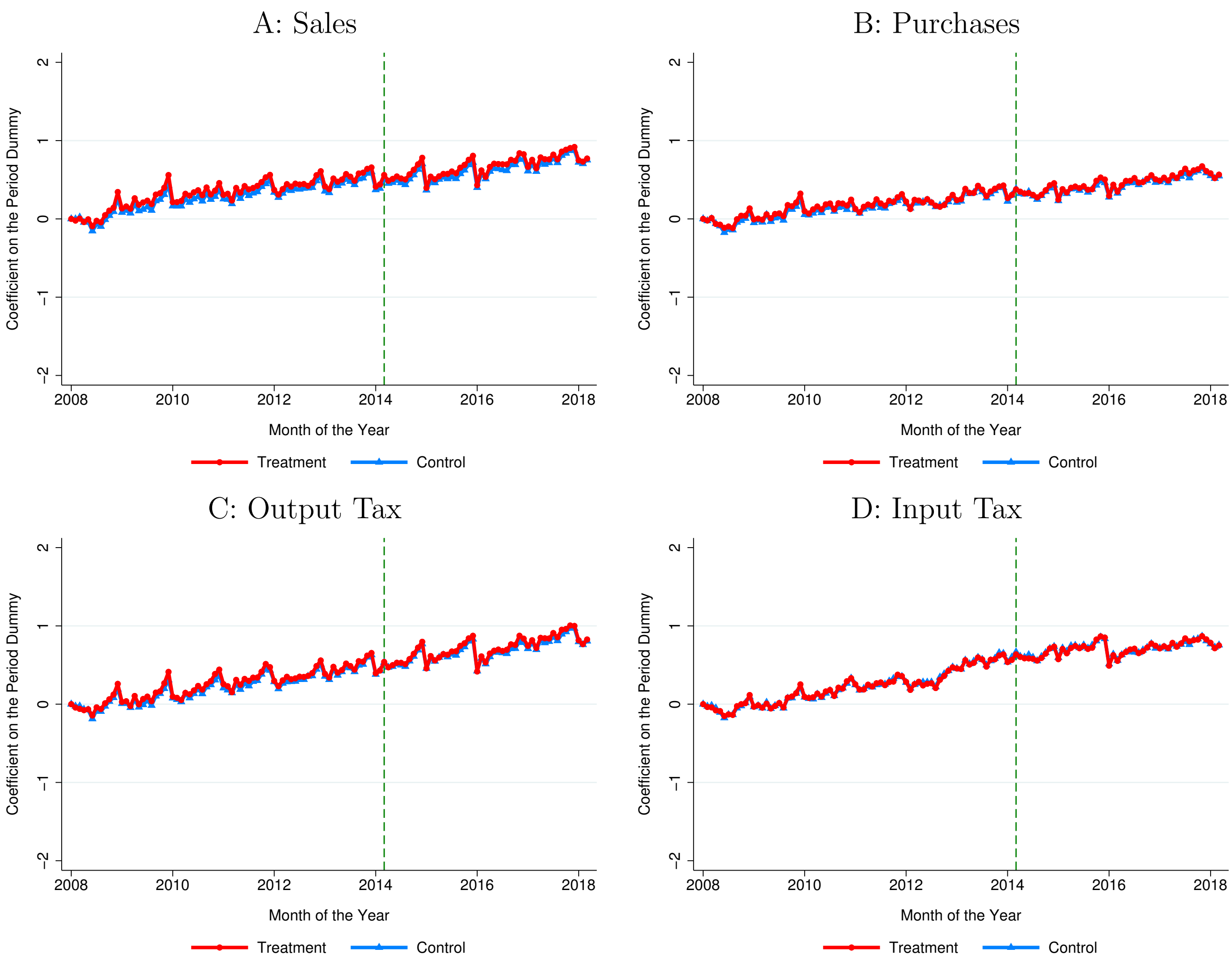
Notes: The figure explores the impacts of audit on future firm behavior. We compare the evolution of four VAT outcomes across the treatment and control groups. The treatment groups consists of firms whose audit was assigned through the first random ballot held on September 13, 2013. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm and month fixed effects, dropping the dummy for July 2008. We then plot the coefficients on the time dummies of these regressions. The sample includes all tax periods from July 2008 to June 2018. The regressions are run separately for the two groups of firms. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines demarcate the date the random computer ballot was held on.

Figure 3.4: Intention to Treat Effects of Audit – First Wave



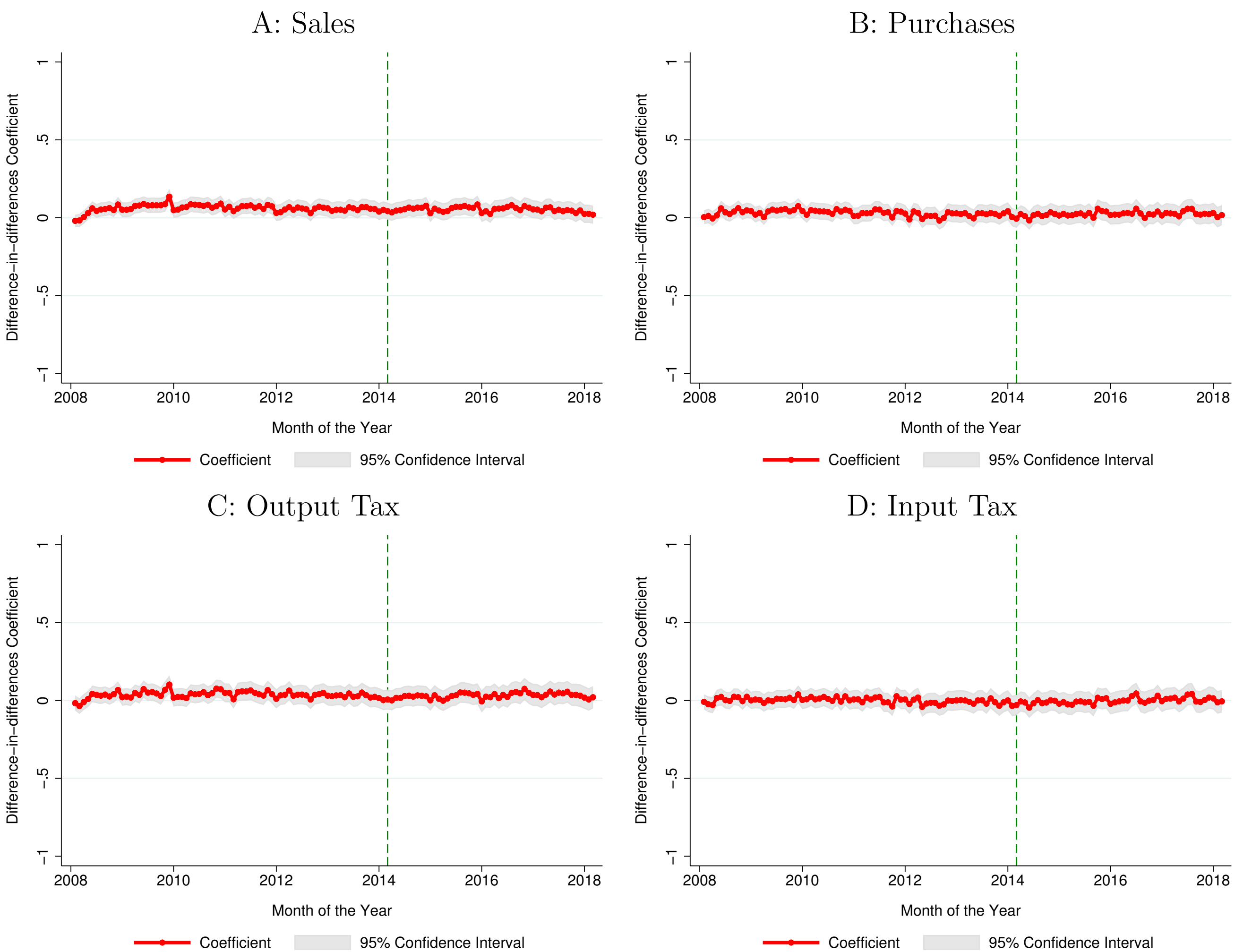
Notes: The figure shows the difference-in-differences version of the plots in Figure 3.1. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm, month, and month×treat dummies, dropping the dummies for July 2008. We then plot the coefficients on the month×treat dummies from these regressions. The gray surface plot shows the 95% confidence interval around the coefficient. The treatment groups consists of firms whose audit was assigned through the first random ballot held on September 13, 2013. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. We cluster standard errors at the firm level. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines demarcate the date the random computer ballot was held on.

Figure 3.5: Intention to Treat Effects of Audit – Second Wave



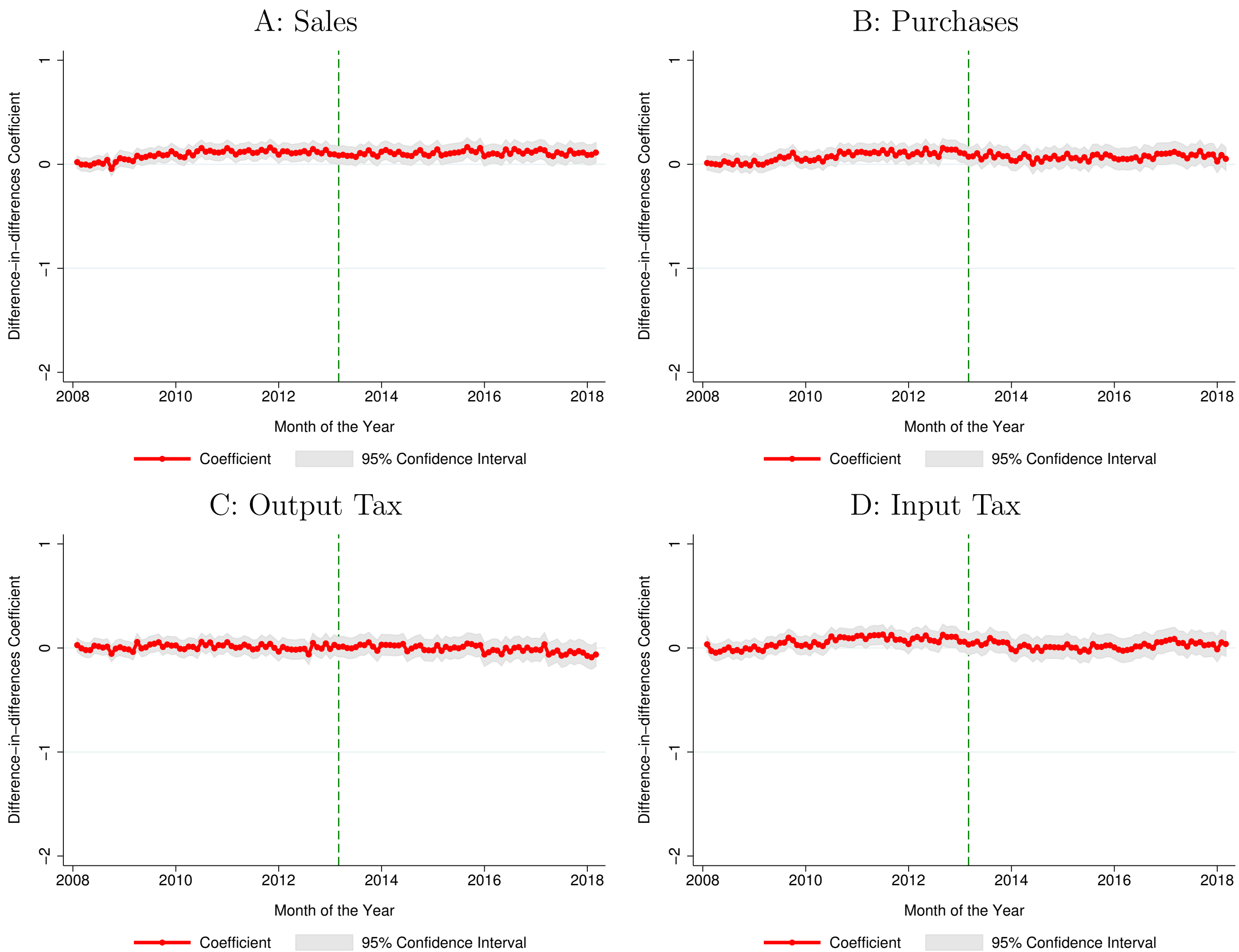
Notes: The figure explores the impacts of audit on future firm behavior. We compare the evolution of four VAT outcomes across the treatment and control groups. The treatment groups consists of firms whose audit was assigned through the first random ballot held on September 25, 2014. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm and month fixed effects, dropping the dummy for July 2008. We then plot the coefficients on the time dummies of these regressions. The sample includes all tax periods from July 2008 to June 2018. The regressions are run separately for the two groups of firms. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines demarcate the date the random computer ballot was held on.

Figure 3.6: Intention to Treat Effects of Audit – Second Wave



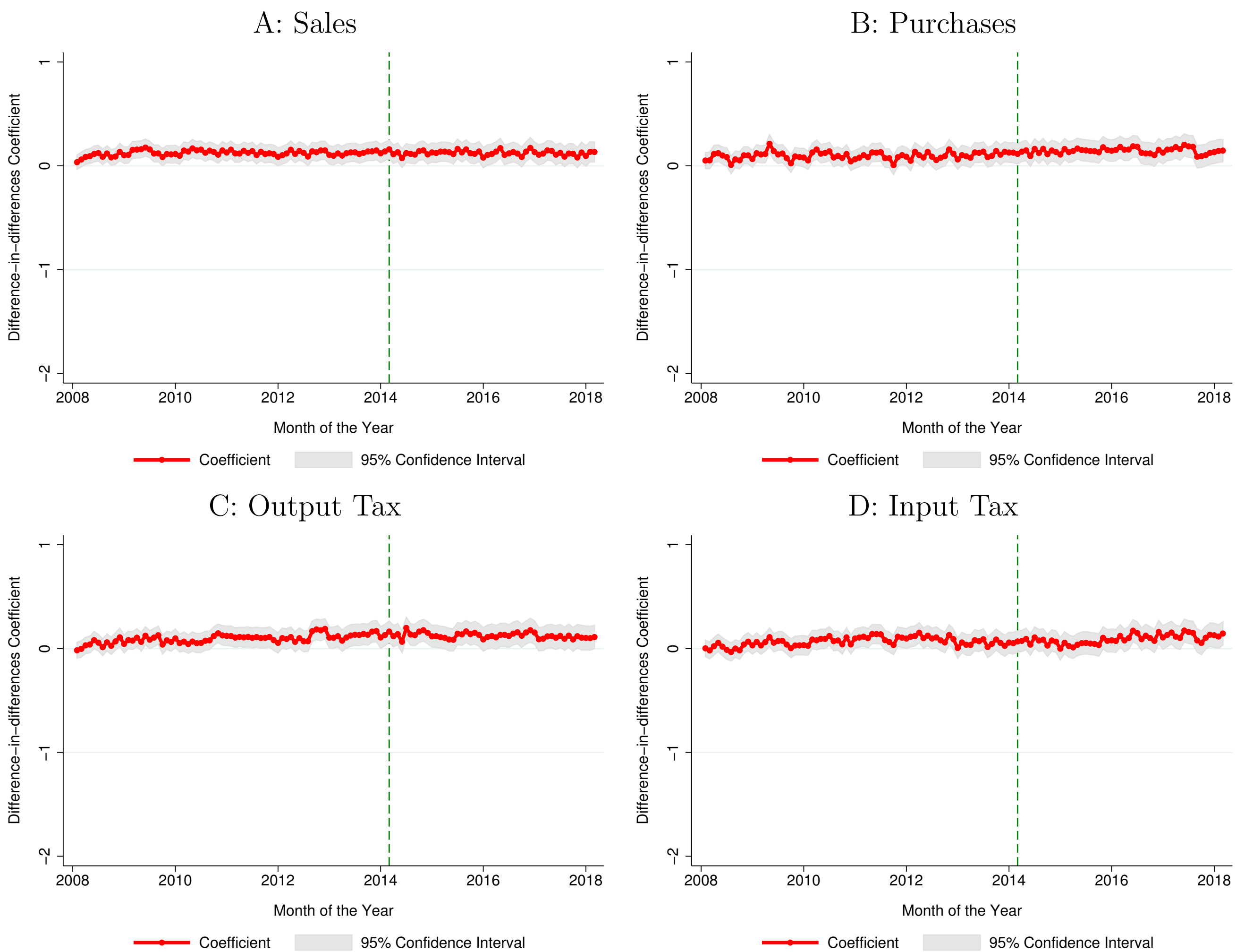
Notes: The figure shows the difference-in-differences version of the plots in Figure 3.5. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm, month, and month×treat dummies, dropping the dummies for July 2008. We then plot the coefficients on the month×treat dummies from these regressions. The gray surface plot shows the 95% confidence interval around the coefficient. The treatment groups consists of firms whose audit was assigned through the first random ballot held on September 25, 2014. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. We cluster standard errors at the firm level. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines demarcate the date the random computer ballot was held on.

Figure 3.7: Audited Vs. Unaudited Firms – First Audit Wave



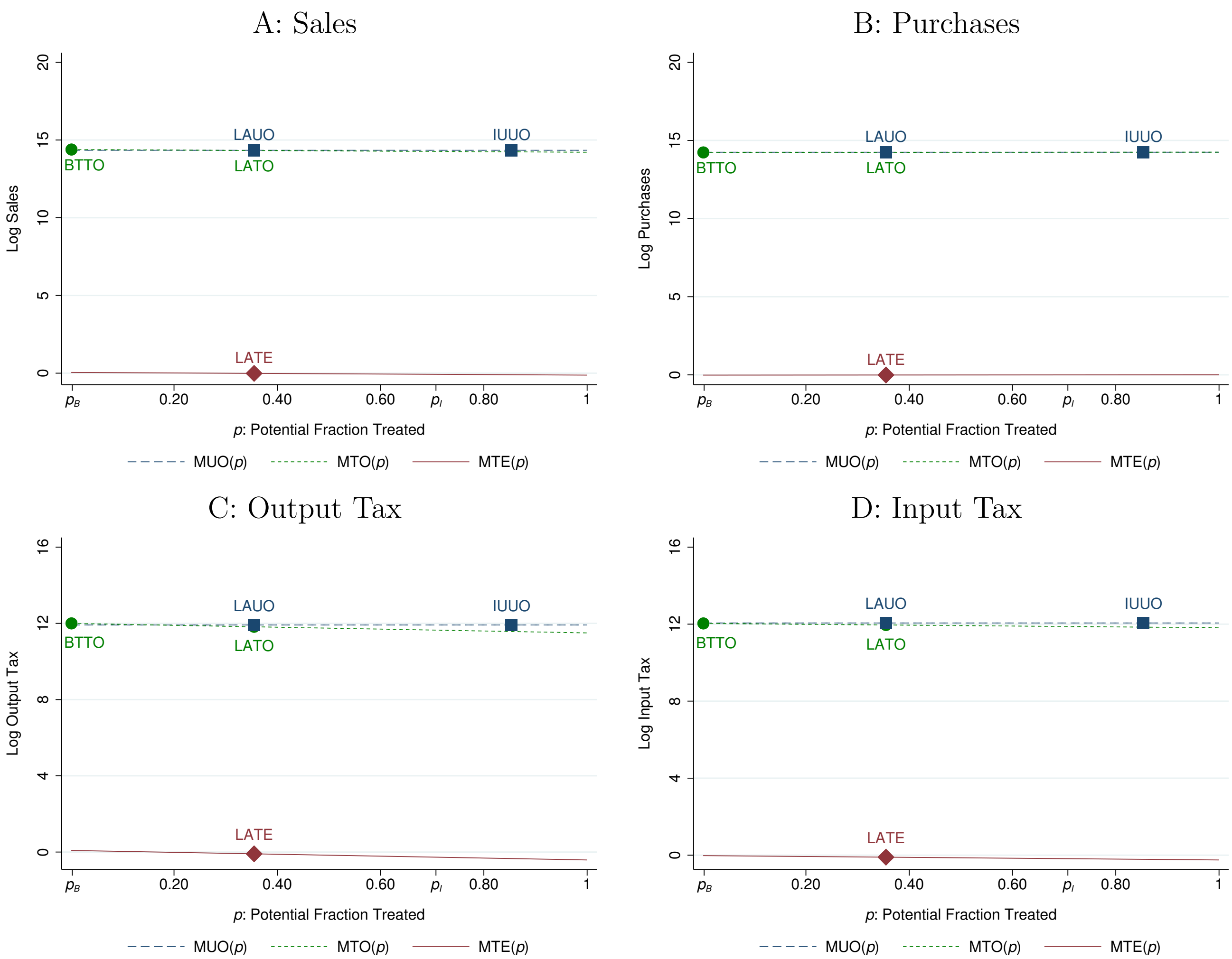
Notes: The figure compares the evolution of outcomes across audited and unaudited firms. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm, month, and month×audit dummies, dropping the dummies for July 2008. We then plot the coefficients on the month×audit dummies from these regressions. The gray surface plot shows the 95% confidence interval around the coefficient. The audit dummy indicates firms whose audit was conducted during the first wave. These includes firms whose audit was assigned through the random computer ballot ($Z_i = 1$) and firms whose audit was initiated by the local tax office on their own accord ($Z_i = 0$). The unaudited firms are all other firms in the population of VAT filers. We cluster standard errors at the firm level. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines denotes September 13, 2013—the date first random computer ballot was held on.

Figure 3.8: Audited Vs. Unaudited Firms – Second Audit Wave



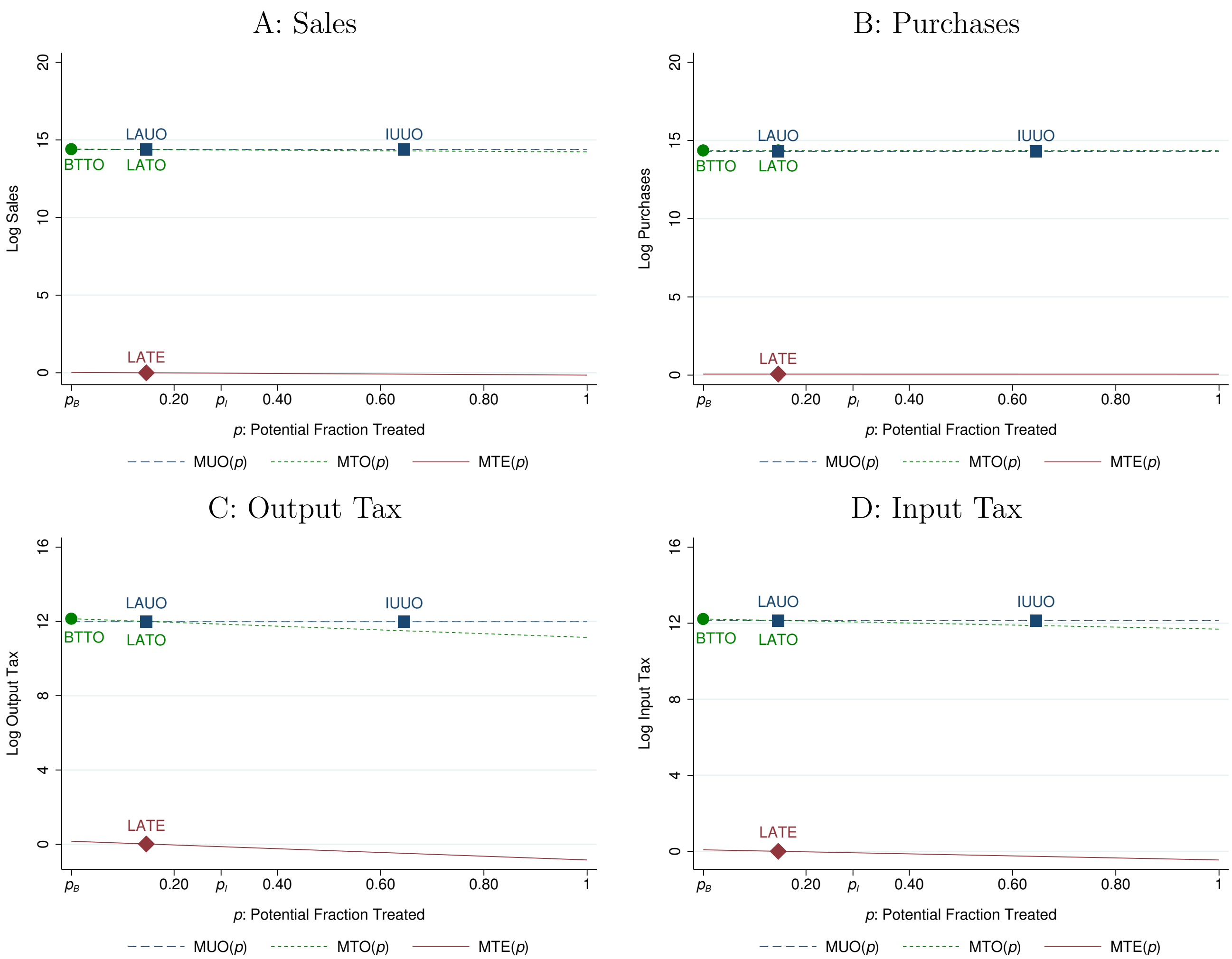
Notes: The figure compares the evolution of outcomes across audited and unaudited firms. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm, month, and month×audit dummies, dropping the dummies for July 2008. We then plot the coefficients on the month×audit dummies from these regressions. The gray surface plot shows the 95% confidence interval around the coefficient. The audit dummy indicates firms whose audit was conducted during the second wave. These includes firms whose audit was assigned through the random computer ballot ($Z_i = 1$) and firms whose audit was initiated by the local tax office on their own accord ($Z_i = 0$). The unaudited firms are all other firms in the population of VAT filers. We cluster standard errors at the firm level. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines denotes September 25, 2014—the date first random computer ballot was held on.

Figure 3.9: Marginal Treatment Effects – First Audit Wave



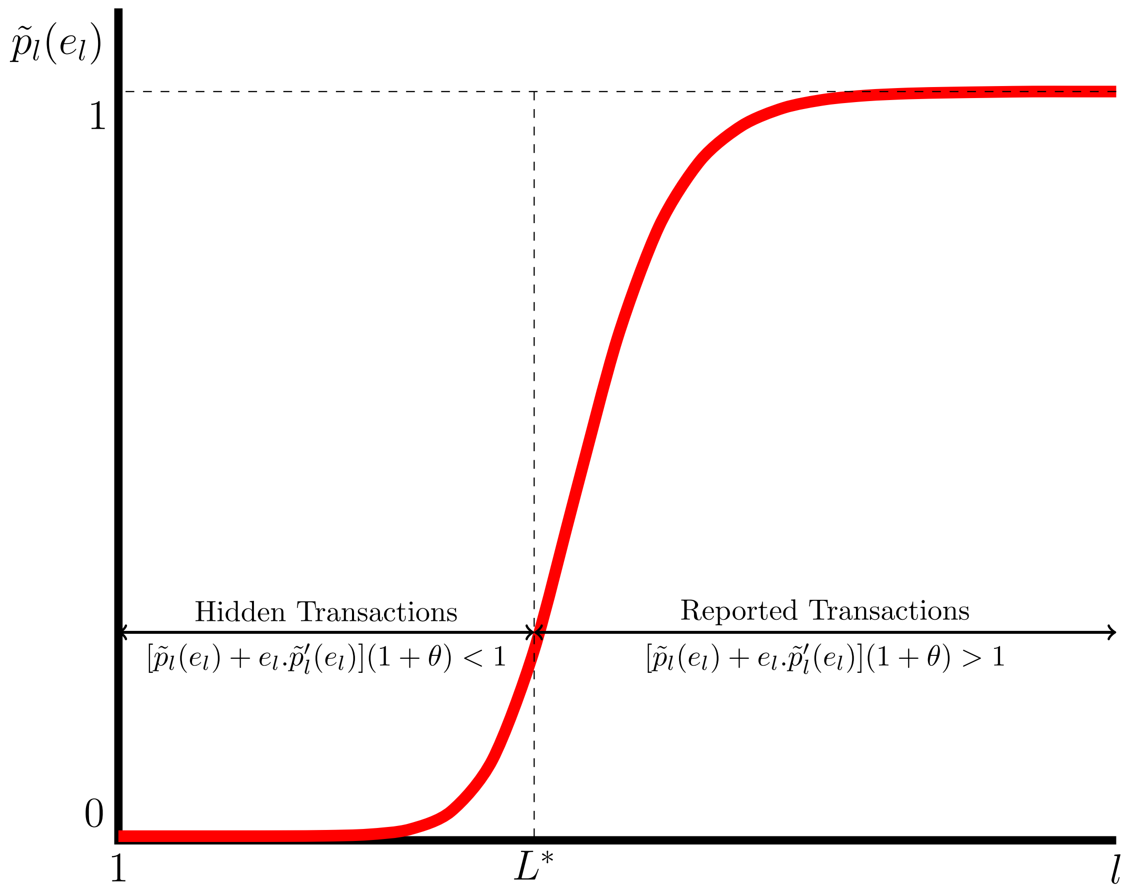
Notes: The figure plots the $MTE(p)$ curve for four outcomes using random assignment in the first audit wave as instrument. Please see Appendix C.2 for technical details. The fraction treated $p \equiv P(D = 1|Z)$ is shown along the horizontal axis. It increases from 0 (no treatment) to 1 (full treatment). We also indicate the baseline treatment probability $p_B \equiv P(D = 1|Z = 0)$ and the intervention treatment probability $p_I \equiv P(D = 1|Z = 1)$ along this axis. The green solid curve shows the marginal treated outcomes curve $MTO(p)$. It is identified at two points indicated in the plot by circular markers. The blue, dashed curve depicts the marginal untreated outcomes curve $MUO(p)$. It is also identified at two points indicated in the plot with square markers. For both curves, we extrapolate between the two points using linearity assumption. The difference between the two curves represents the $MTE(p)$. Since in our setting all three curves sit above each other, we lift both $MTO(p)$ and $MUO(p)$ up by adding the constant from the corresponding regression to distinguish them from the primary object of our interest $MTE(p)$.

Figure 3.10: Marginal Treatment Effects – Second Audit Wave



Notes: The figure plots the $MTE(p)$ curve for four outcomes using random assignment in the second audit wave as instrument. Please see Appendix C.2 for technical details. The fraction treated $p \equiv P(D = 1|Z)$ is shown along the horizontal axis. It increases from 0 (no treatment) to 1 (full treatment). We also indicate the baseline treatment probability $p_B \equiv P(D = 1|Z = 0)$ and the intervention treatment probability $p_I \equiv P(D = 1|Z = 1)$ along this axis. The green solid curve shows the marginal treated outcomes curve $MTO(p)$. It is identified at two points indicated in the plot by circular markers. The blue, dashed curve depicts the marginal untreated outcomes curve $MUO(p)$. It is also identified at two points indicated in the plot with square markers. For both curves, we extrapolate between the two points using linearity assumption. The difference between the two curves represents the $MTE(p)$. Since in our setting all three curves sit above each other, we lift both $MTO(p)$ and $MUO(p)$ up by adding the constant from the corresponding regression to distinguish them from the primary object of our interest $MTE(p)$.

Figure 3.11: Probability of Detection



Notes: The figure plots the probability of detection faced by a typical firm. We arrange L transactions carried out by the firm in term of the detection probability they entail $p_l(e_l)$ in ascending order. The probability of transaction is low if the other party to the transaction is (1) a consumer, (2) an unregistered firm, or (3) a firm willing to collude. In all these case, the transaction does not create any third-party information for the government. The probability of detection is high otherwise. The curve accordingly turns sharply once transactions between arm-length parties unwilling to collude begin. The transaction L^*represents the first transaction for which the detection probability is so high that inequality (3.12) fails. The firm would accordingly report transactions $[L^*, L]$, hiding the rest. Note that the threshold L^*would vary across firms depending among other things on their size, industry, and trading network.

Table 3.1: Descriptive Statistics of Audit I

Audit	Tax	Ballot	Audits Assigned			Audits Conducted	
Wave	Year	Date	Mode	Number	Percent	Assigned	Unassigned
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	2013	Sep 13, 2013	Random	4,926	5%	3,482	521
2	2014	Sep 25, 2014	Random	12,447	12%	3,612	293
3	2015	Sep 14, 2015	Random	8,372	7.5%	1,122	164
4	2016	Jan 05, 2017	Parametric	8,935	7.5%	884	332
5	2017	Apr 12, 2018	Parametric	8,785	7.5%	852	352

Notes*:* The table reports some descriptive statistics of the five audit waves in our sample. Column (2) reports the tax year during which the computer ballot to draw audit cases was held. Column (3) reports the exact ballot date. The ballot was random for the first three waves and parametric for the next two. The volume of cases picked during the ballot is mentioned in Column (5) in numbers and in Column (6) as the proportion of population. Column (7) reports the number of audits completed out of those assigned through the computer ballot. Column (8), on the other hand, reports the number of audits initiated by the local tax office on their own accord. During the five audit waves, a total of 43,625 cases were picked for audit through computer ballots. Out of these, the tax identifiers of 218 were inaccurate. We were therefore unable to merge these 218 cases with VAT and audit records. We accordingly drop these 218 cases from the sample and focus instead on the 43,465 audits assigned through the computer ballot as reported in Column (5).

Table 3.2: Descriptive Statistics of Audit II

Audit	Audits Initiated			Amount Detected		
Wave	Within 1 Month	Within 3 Months	Within 6 Months	Mean	Median	90th Percentile
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	0.646	0.942	0.950	617	0	165
2	0.925	0.993	0.998	619	0	100
3	0.852	0.945	0.964	4,098	0	158

Notes*:* The table presents a few descriptive statistics of randomly assigned audits during the first three audit waves. Columns (2)-(4) report the time lag between the assignment and initiation of audit. Column (2), for example, shows that around 65% of audits assigned in the first random ballot were initiated with the first month of assignment. This ratio was 93% and 85% for the next two audit waves. Columns (5)-(7) report the amount detected during each wave of audit. Column (5) reports the mean amount detected in PKR thousands. The US$-PKR exchange rate during this time (2013) was around 100. The next columns of the table report the median and the 90th percentile of the amount detected, illustrating that it is highly skewed toward right with the mean significantly larger than the median for all three audit waves.

Table 3.3: Randomization Test

	First Wave				Second Wave				Third Wave			
	Mean ($Z_i = 0$)	Mean ($Z_i = 1$)	Diff. in Means	SE	Mean ($Z_i = 0$)	Mean ($Z_i = 1$)	Diff. in Means	SE	Mean ($Z_i = 0$)	Mean ($Z_i = 1$)	Diff. in Means	SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A: VAT Outcomes												
1. Sales	14.251	14.282	0.031	0.043	14.278	14.298	0.020	0.026	14.335	14.831	0.496	0.029
2. Purchases	14.081	14.095	0.014	0.047	14.234	14.186	-0.048	0.029	14.264	14.248	-0.015	0.035
3. Output Tax	11.671	11.707	0.036	0.049	11.791	11.768	-0.024	0.030	11.969	11.953	-0.017	0.035
4. Input Tax	11.768	11.802	0.033	0.052	11.990	11.911	-0.079	0.031	12.149	11.886	-0.263	0.037
5. Tax Payable	10.200	10.300	0.100	0.063	10.392	10.360	-0.032	0.041	10.570	10.830	0.260	0.045
6. Tax Paid	9.532	9.607	0.076	0.058	9.805	9.785	-0.020	0.034	9.850	10.338	0.488	0.039
7. Exports	15.288	15.169	-0.119	0.114	14.904	15.145	0.241	0.068	14.619	15.655	1.036	0.064
8. Imports	14.905	14.887	-0.018	0.078	14.858	14.843	-0.015	0.048	14.878	15.902	1.024	0.076
9. Refund	12.037	11.884	-0.153	0.152	12.214	12.188	-0.026	0.089	12.086	12.424	0.338	0.093
10. Carry Forward	11.642	11.667	0.026	0.078	12.010	12.160	0.150	0.046	12.162	12.248	0.086	0.050
B: Firm Characteristics												
11. Manufacturer	0.339	0.350	0.010	0.010	0.314	0.339	0.025	0.006	0.215	0.786	0.572	0.006
12. Importer	0.111	0.109	-0.003	0.006	0.124	0.118	-0.006	0.004	0.159	0.019	-0.140	0.002
13. Exporter	0.025	0.019	-0.005	0.003	0.040	0.025	-0.016	0.002	0.050	0.021	-0.029	0.002
14. Distributor	0.028	0.030	0.001	0.003	0.031	0.034	0.003	0.002	0.036	0.011	-0.025	0.002
15. Wholesaler	0.240	0.241	0.001	0.008	0.229	0.240	0.011	0.005	0.251	0.046	-0.205	0.003
16. Service Provider	0.193	0.192	-0.002	0.008	0.193	0.185	-0.009	0.005	0.208	0.099	-0.110	0.005
17. Major City	0.640	0.636	-0.004	0.010	0.631	0.639	0.008	0.006	0.625	0.650	0.024	0.007
18. LTU	0.013	0.013	0.000	0.004	0.012	0.008	-0.004	0.002	0.005	0.042	0.037	0.003
19. Years Registered	12.987	13.680	0.694	0.109	11.745	12.967	1.222	0.070	10.496	13.607	3.111	0.091
20. Textile	0.162	0.163	0.001	0.008	0.143	0.152	0.009	0.005	0.108	0.266	0.157	0.006

Notes*:* The table runs balance tests on the three randomization waves in our sample. For each outcome, we estimate model (3.9) restricting the sample to the baseline period only. The baseline period is June 2012 for the first, June 2013 for the second, and June 2014 for the third randomization wave. The last two columns for each randomization wave report the coefficient $\hat{\beta}$ and its standard error from the model. The details of the variables used here are provided in Appendix C.1.

Table 3.4: Audit Findings

	# Audits	Sales	Amount Detected		VAT Paid at the Baseline		Evasion Rate
			PKR	% of Sales	PKR	% of Sales	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: First Audit Wave							
All Audited Firms	3,482	498.4	2.15	0.43	28.16	5.65	7.1
Amount Detected > 0	986	137.0	2.15	1.57	3.20	2.33	40.2
Size Quartile 1	1,057	0.0	0.06	684.76	0.00	8.78	98.7
Size Quartile 2	824	1.7	0.07	3.94	0.04	2.52	61.0
Size Quartile 3	809	12.3	0.22	1.75	0.21	1.67	51.1
Size Quartile 4	792	484.3	1.80	0.37	27.91	5.76	6.1
B: Second Audit Wave							
All Audited Firms	3,612	2200.0	2.24	0.10	88.37	4.02	2.5
Amount Detected > 0	1,220	264.6	2.24	0.84	7.52	2.84	22.9
Size Quartile 1	1,007	0.4	0.04	10.21	0.02	3.81	72.8
Size Quartile 2	892	4.9	0.17	3.37	0.11	2.15	61.0
Size Quartile 3	862	24.4	0.22	0.89	0.30	1.24	41.8
Size Quartile 4	851	2170.2	1.81	0.08	87.95	4.05	2.0

Notes: The table presents descriptive statistics of audit outcomes. The first column reports the number of audits conducted for each group of firms indicated in the corresponding row. Aggregate turnover of this group for the baseline year in PKR billions is reported in the next column. The next two columns report the amount detected by audit, in PKR billions in column 3 and as a percent of aggregate sales in column 4. Columns 5-6 report the VAT paid at the baseline by the group of firms indicated in the corresponding row, in PKR billions in column 5 and as a percent of aggregate sales in column 6. The last column presents the evasion rate implied by the detected amount. It is calculated as the ratio of columns 4 and 6 (alternatively columns 3 and 5).

Table 3.5: Impact of Audit on Firm Behavior – First Wave

	Impacts After One Year					Impacts After Three Years				
	Sales	Purchases	Output Tax	Input Tax	Tax Payable	Sales	Purchases	Output Tax	Input Tax	Tax Payable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A: ITT Estimates										
assign × after	-0.010	-0.010	-0.016	-0.017	-0.037	-0.007	-0.021	-0.025	-0.036	-0.016
	(0.016)	(0.018)	(0.021)	(0.023)	(0.028)	(0.017)	(0.019)	(0.022)	(0.023)	(0.028)
Observations	2,831,140	2,468,502	2,086,889	2,099,210	1,415,795	3,839,502	3,328,628	2,884,225	2,906,045	1,913,096
B: LATE Estimates										
audit × after	-0.014	-0.014	-0.023	-0.024	-0.051	-0.010	-0.030	-0.035	-0.051	-0.022
	(0.023)	(0.027)	(0.030)	(0.033)	(0.039)	(0.024)	(0.028)	(0.031)	(0.033)	(0.039)
Observations	2,831,140	2,468,502	2,086,889	2,099,210	1,415,795	3,839,502	3,328,628	2,884,225	2,906,045	1,913,096
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table estimates the impact of audit on firms' future behavior. In the top panel, the coefficient assign × after shows $\hat{\gamma}$ from model (3.10), where the dummy variable assign$_i$ denotes that firm i's audit was assigned through the first random ballot held on September 13, 2013. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable after$_t$ indicates that month t falls after the date of the ballot. The sample includes periods up to October 2014 for the first five columns and periods up to October 2016 for the rest. Panel B shows the corresponding results from 2sls regressions, where the endogenous variable audit$_i$ is instrumented by the initial random assignment. Standard errors are in parenthesis, which have been clustered at the firm level.

Table 3.6: Impact of Audit on Firm Behavior – Second Wave

	Impacts After One Year					Impacts After Three Years				
	Sales	Purchases	Output Tax	Input Tax	Tax Payable	Sales	Purchases	Output Tax	Input Tax	Tax Payable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A: ITT Estimates										
assign × after	-0.021	-0.021	-0.030	-0.026	-0.022	-0.010	-0.009	-0.013	-0.007	0.006
	(0.010)	(0.012)	(0.012)	(0.013)	(0.016)	(0.010)	(0.012)	(0.013)	(0.013)	(0.016)
Observations	3,133,061	2,725,243	2,343,583	2,357,343	1,568,363	4,159,404	3,587,740	3,088,403	3,137,794	2,034,932
B: LATE Estimates										
audit × after	-0.071	-0.073	-0.109	-0.091	-0.081	-0.032	-0.033	-0.044	-0.025	0.022
	(0.033)	(0.043)	(0.043)	(0.046)	(0.058)	(0.035)	(0.043)	(0.044)	(0.045)	(0.057)
Observations	3,133,061	2,725,243	2,343,583	2,357,343	1,568,363	4,159,404	3,587,740	3,088,403	3,137,794	2,034,932
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table estimates the impact of audit on firms' future behavior. In the top panel, the coefficient assign × after shows $\hat{\gamma}$ from model (3.10), where the dummy variable assign_i denotes that firm i's audit was assigned through the first random ballot held on September 25, 2014. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable after_t indicates that month t falls after the date of the ballot. The sample includes periods up to October 2015 for the first five columns and periods up to October 2017 for the rest. Panel B shows the corresponding results from 2sls regressions, where the endogenous variable audit_i is instrumented by the initial random assignment. Standard errors are in parenthesis, which have been clustered at the firm level.

Table 3.7: Impacts of Random Audits Assigned in the First Wave – Other Outcomes

	Impacts After One Year					Impacts After Three Years				
	Exports	Imports	Tax Paid	Refund	Carry Forward	Exports	Imports	Tax Paid	Refund	Carry Forward
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<u>A: ITT Estimates</u>										
assign × after	0.013	0.047	-0.052	-0.116	-0.049	0.027	0.035	-0.025	-0.070	-0.085
	(0.037)	(0.028)	(0.031)	(0.092)	(0.040)	(0.038)	(0.027)	(0.033)	(0.091)	(0.040)
Observations	317,130	570,949	1,161,513	234,207	1,594,740	450,661	838,590	1,723,448	287,241	2,490,894
<u>B: LATE Estimates</u>										
audit × after	0.018	0.073	-0.072	-0.175	-0.071	0.037	0.054	-0.035	-0.102	-0.124
	(0.051)	(0.043)	(0.043)	(0.138)	(0.058)	(0.053)	(0.042)	(0.046)	(0.134)	(0.059)
Observations	317,130	570,949	1,161,513	234,207	1,594,740	450,661	838,590	1,723,448	287,241	2,490,894
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table estimates the impact of audit on firms' future behavior. In the top panel, the coefficient assign × after shows $\hat{\gamma}$ from model (3.10), where the dummy variable assign$_i$ denotes that firm i's audit was assigned through the first random ballot held on September 13, 2013. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable after$_t$ indicates that month t falls after the date of the ballot. The sample includes periods up to October 2014 for the first five columns and periods up to October 2016 for the rest. Panel B shows the corresponding results from 2sls regressions, where the endogenous variable audit$_i$ is instrumented by the initial random assignment. Standard errors are in parenthesis, which have been clustered at the firm level.

Table 3.8: Extensive Margin Impact of Random Audits

Outcome:	$1(Return\ Filed_{it})$					
Random Draw Held On:	September 13, 2013		September 25, 2014		September 14, 2015	
Impacts After:	One Year	Three Years	One Year	Three Years	One Year	Three Years
	(1)	(2)	(3)	(4)	(5)	(6)
<u>A: ITT Estimates</u>						
assign × after	0.002	0.004	0.008	0.009	0.010	0.008
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	7,097,120	9,852,941	8,129,498	11,062,795	8,502,891	11,171,180
<u>B: LATE Estimates</u>						
audit × after	0.002	0.006	0.027	0.029	0.075	0.058
	(0.002)	(0.003)	(0.004)	(0.004)	(0.010)	(0.009)
Observations	7,097,120	9,852,941	8,129,498	11,062,795	8,502,891	11,171,180
Mean of the Dependent Variable	0.955	0.955	0.956	0.956	0.956	0.956
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table estimates the impact of audit on firms' extensive margin behavior. We estimate model (3.10) using an indicator that the firm filed its VAT return for the period (month) t as the outcome variable. In the top panel, the coefficient assign × after shows $\hat{\gamma}$ from the model. The dummy variable assign$_i$ denotes that firm i's audit was assigned through the random ballot held on the date indicated in the heading of each column. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable after$_t$ indicates that month t falls after the date of the ballot. The sample for odd columns includes periods up to one year after the ballot and for even columns up to three years after the ballot. Panel B shows the corresponding results from 2sls regressions, where the endogenous variable audit$_i$ is instrumented by the initial random assignment. Standard errors are in parenthesis, which have been clustered at the firm level.

Table 3.9: Selection in Compliance? Audited Vs. Non-Audited Firms

	First Wave				Second Wave			
	Mean ($D_i = 0$)	Mean ($D_i = 1$)	Difference in Means	Standard Error	Mean ($D_i = 0$)	Mean ($D_i = 1$)	Difference in Means	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>A: VAT Outcomes</u>								
1. Sales	14.547	14.816	0.269	0.044	14.553	14.776	0.222	0.043
2. Purchases	14.311	14.567	0.255	0.048	14.438	14.519	0.080	0.049
3. Output Tax	11.936	12.214	0.279	0.050	12.031	12.199	0.168	0.051
4. Input Tax	12.006	12.328	0.322	0.051	12.196	12.230	0.033	0.051
5. Tax Payable	10.537	10.866	0.328	0.068	10.698	10.870	0.172	0.075
6. Tax Paid	10.039	10.435	0.397	0.062	10.221	10.359	0.137	0.063
7. Exports	15.752	15.705	-0.047	0.114	15.353	15.793	0.440	0.096
8. Imports	15.183	15.261	0.078	0.075	15.096	15.235	0.139	0.074
9. Refund	12.410	12.673	0.263	0.139	12.578	12.667	0.089	0.130
10. Carry Forward	11.926	12.192	0.266	0.081	12.276	12.446	0.170	0.083
<u>B: Firm Characteristics</u>								
11. Manufacturer	0.383	0.448	0.064	0.010	0.361	0.418	0.056	0.009
12. Importer	0.105	0.087	-0.018	0.006	0.116	0.111	-0.005	0.006
13. Exporter	0.023	0.016	-0.007	0.003	0.036	0.013	-0.023	0.003
14. Distributor	0.027	0.026	-0.001	0.004	0.029	0.028	-0.001	0.004
15. Wholesaler	0.214	0.196	-0.018	0.008	0.206	0.219	0.012	0.008
16. Service Provider	0.190	0.174	-0.016	0.008	0.189	0.166	-0.023	0.008
17. Major City	0.661	0.661	0.000	0.000	0.654	0.654	0.000	0.000
18. LTU	0.045	0.045	-0.000	0.000	0.039	0.039	-0.000	0.000
19. Years Registered	13.499	14.729	1.230	0.117	12.388	14.221	1.833	0.119
20. Textile	0.165	0.171	0.005	0.007	0.148	0.160	0.012	0.006

Notes*:* The table explores selection in audit, comparing audited and unaudited firms. We estimate a version of model (3.9), regressing the outcome in each row on two dummy variables (D_i and $corporate_i$) and tax office fixed effects. We restrict the sample to the baseline period only. The dummy variable D_i takes the value 1 for all audited firms including those whose audit was assigned through the random ballot and those whose audit was taken up by the local tax office of its own accord. The unaudited firms ($D_i = 0$) include all other firms in the eligible sample. The baseline period is June 2012 for the first and June 2013 for the second audit wave. The last two columns for each wave report the coefficient $\hat{\beta}$ and its standard error from the model. The details of the variables used here are provided in Appendix C.1.

Table 3.10: Selection in Compliance? Audited Vs. Non-Audited Firms (Within $Z_i = 1$ Group)

	2013 Draw				2014 Draw			
	Mean ($D_i = 0$)	Mean ($D_i = 1$)	Difference in Means	Standard Error	Mean ($D_i = 0$)	Mean ($D_i = 1$)	Difference in Means	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>A: VAT Outcomes</u>								
1. Sales	14.567	14.569	0.001	0.095	14.560	14.633	0.073	0.061
2. Purchases	14.360	14.312	-0.048	0.108	14.393	14.410	0.017	0.066
3. Output Tax	11.885	12.005	0.120	0.104	11.982	12.094	0.112	0.069
4. Input Tax	11.944	12.075	0.131	0.117	12.131	12.115	-0.017	0.070
5. Tax Payable	10.524	10.666	0.142	0.132	10.648	10.715	0.067	0.101
6. Tax Paid	9.935	10.175	0.240	0.129	10.206	10.173	-0.033	0.083
7. Exports	15.602	15.678	0.076	0.285	15.476	15.897	0.422	0.226
8. Imports	15.131	15.150	0.018	0.178	15.057	15.154	0.097	0.101
9. Refund	11.650	12.482	0.832	0.385	12.502	12.681	0.179	0.257
10. Carry Forward	11.833	12.023	0.190	0.173	12.424	12.331	-0.093	0.108
<u>B: Firm Characteristics</u>								
11. Manufacturer	0.364	0.406	0.042	0.022	0.378	0.397	0.019	0.013
12. Importer	0.115	0.096	-0.019	0.016	0.107	0.120	0.013	0.010
13. Exporter	0.018	0.017	-0.001	0.006	0.022	0.022	0.000	0.004
14. Distributor	0.030	0.027	-0.003	0.008	0.033	0.029	-0.003	0.005
15. Wholesaler	0.228	0.210	-0.018	0.020	0.218	0.215	-0.003	0.012
16. Service Provider	0.186	0.188	0.001	0.017	0.185	0.170	-0.015	0.011
17. Major City	0.655	0.655	-0.000	0.000	0.659	0.659	-0.000	0.000
18. LTU	0.043	0.043	0.000	0.000	0.035	0.035	0.000	0.000
19. Years Registered	13.865	14.313	0.448	0.258	13.175	14.222	1.047	0.167
20. Textile	0.163	0.167	0.005	0.017	0.158	0.154	-0.004	0.009

Notes*:* The table explores selection in audit, comparing audited and unaudited firms within the sample drawn for audit in the corresponding random ballot. We estimate a version of model (3.9), regressing the outcome in each row on two dummy variables (D_i and $corporate_i$) and tax office fixed effects. We restrict the sample to the baseline period only. The dummy variable D_i takes the value 1 for firms whose audit was conducted. The unaudited firms ($D_i = 0$) include all other firms in the randomly drawn sample. The baseline period is June 2012 for the first and June 2013 for the second audit wave. The last two columns for each wave report the coefficient $\hat{\beta}$ and its standard error from the model. The details of the variables used here are provided in Appendix C.1.

Table 3.11: Selection in Compliance? Audited Vs. Non-Audited Firms (Within $Z_i = 0$ Group)

	2013 Draw				2014 Draw			
	Mean ($D = 0$)	Mean ($D = 1$)	Difference in Means	Standard Error	Mean ($D = 0$)	Mean ($D = 1$)	Difference in Means	Standard Error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>A: VAT Outcomes</u>								
1. Sales	14.548	15.693	1.145	0.086	14.556	15.776	1.220	0.149
2. Purchases	14.312	15.330	1.018	0.095	14.446	15.549	1.103	0.169
3. Output Tax	11.936	12.958	1.022	0.098	12.040	13.046	1.006	0.170
4. Input Tax	12.008	13.045	1.037	0.097	12.208	13.053	0.844	0.171
5. Tax Payable	10.537	11.704	1.167	0.148	10.708	11.979	1.270	0.235
6. Tax Paid	10.040	11.220	1.180	0.132	10.227	11.390	1.163	0.190
7. Exports	15.750	16.009	0.258	0.216	15.330	16.019	0.689	0.372
8. Imports	15.183	15.473	0.290	0.129	15.101	15.689	0.588	0.198
9. Refund	12.425	13.168	0.743	0.291	12.585	13.168	0.583	0.381
10. Carry Forward	11.927	12.857	0.930	0.164	12.255	13.138	0.883	0.279
<u>B: Firm Characteristics</u>								
11. Manufacturer	0.384	0.622	0.239	0.021	0.359	0.531	0.172	0.030
12. Importer	0.105	0.049	-0.055	0.010	0.117	0.110	-0.007	0.019
13. Exporter	0.023	0.017	-0.006	0.004	0.037	0.026	-0.011	0.002
14. Distributor	0.026	0.020	-0.006	0.007	0.028	0.017	-0.012	0.014
15. Wholesaler	0.214	0.133	-0.081	0.012	0.205	0.182	-0.023	0.022
16. Service Provider	0.190	0.111	-0.079	0.016	0.190	0.103	-0.087	0.025
17. Major City	0.661	0.661	0.000	0.000	0.652	0.652	-0.000	0.000
18. LTU	0.045	0.045	-0.000	0.000	0.040	0.040	0.000	0.000
19. Years Registered	13.493	16.379	2.886	0.288	12.266	14.597	2.331	0.437
20. Textile	0.165	0.187	0.021	0.017	0.147	0.208	0.061	0.025

Notes: The table explores selection in audit, comparing audited and unaudited firms excluding from the sample firms drawn for audit in the corresponding random ballot. We estimate a version of model (3.9), regressing the outcome in each row on two dummy variables (D_i and $corporate_i$) and tax office fixed effects. We restrict the sample to the baseline period only. The dummy variable D_i takes the value 1 for firms whose audit was conducted. The baseline period is June 2012 for the first and June 2013 for the second audit wave. The last two columns for each wave report the coefficient $\hat{\beta}$ and its standard error from the model. The details of the variables used here are provided in Appendix C.1.

Chapter 4

Conclusion

This dissertation analyzes evasion and enforcement issues in VAT with special reference to developing countries using three different policy reforms in Pakistan. The first essay provides empirical evidence on the effectiveness of computerization and measures its impact on input tax credits claimed on domestic purchases. Conversely, it also shows that evasion and fraudulent transaction are rampant absent an extensive enforcement mechanism which can utilize third party information effectively. The second essay studies a minimum tax regime within a standard VAT system. The results of this study show that commercial importers may use informal sector to lower their declared value at the cost of revenue. Tax authorities, thus, may be forced to impose a minimum tax burden on these evasion prone firms. Third chapter examines a randomized audit program and its impact on subsequent tax filing and revenue reporting behavior of selected and audited firms. The results show that there is no effect of audit and detection on post audit behavior of firms.

In the first chapter, I study the introduction of a risk-based evaluation system which cross checks input tax credits and automatically invalidates suspicious and fraudulent invoices. Using a difference-in-differences empirical estimation, I find that input tax credits for domestically operating firms (treatment) went down by 50% on average. The response depended on firm's structure and category of operations. Manufacturers showed an average decline of 30% whereas for non-manufacturers it was 70%. Input tax credits for Sole Proprietors, Partnerships and Companies decreased by 70%, 60% and 30% respectively. Computerized enforcement was very effective, and it was able to check fake credit invoices. It also indicates that absent this effective enforcement, a huge amount of revenue can be evaded in developing countries who have limited enforcement capacity. I argue that implementation of VAT in developing countries was premature because of enforcement capacity constraints.

The second essay deals with minimum taxes within VAT. I utilize kinks produced by minimum value addition tax imposed on commercial importers for studying tax evasion on sales made to informal sector. Pakistan imposes a minimum value addition tax on importers at the time of import which is adjustable only when the declared value addition at post importation stage is more than minimum already paid. I find that importers bunch strongly around this minimum value addition threshold and utilizing changes in threshold over time I can estimate evasion. Depending on the threshold, fifty percent firms, on average, bunch at or near the bunch point. I estimate elasticity of tax remitted at 0.22 which implies that 78% of tax would be evaded in this setting. This response implies that informal sector is not sufficiently burdened. It strengthens the argument that replacement of high tariffs at import with revenue neutral VAT would decrease welfare.

The third essay is based on a randomized audit program which was initiated by FBR in 2013 instead of traditional parametric audit. The change in audit selection method should result in different responses. First, firms selected for audits are expected to change their behavior realizing that they are now less (more) likely to be audited. But firms show no change in behavior over time across any of the key parameters. The result is surprising because, theoretically, the change in probability of audit should affect firm behavior.

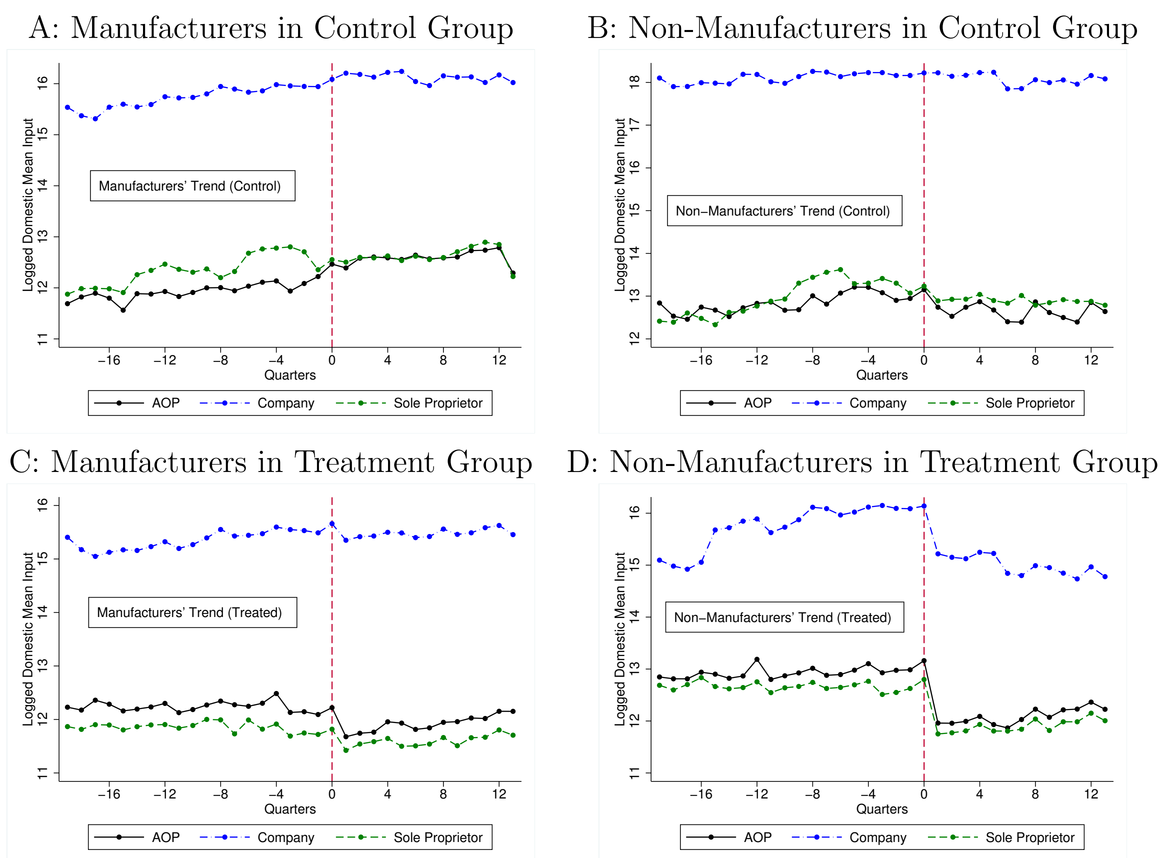
In conclusion, these essays show that firms behave differently in limited enforcement capacity regimes. The volume of evasion is high and traditional instruments to increase compliance such as audit do not have predicted effect in these settings.

Appendix A

Appendix to Chapter 1

A.1 Additional Tables and Figures

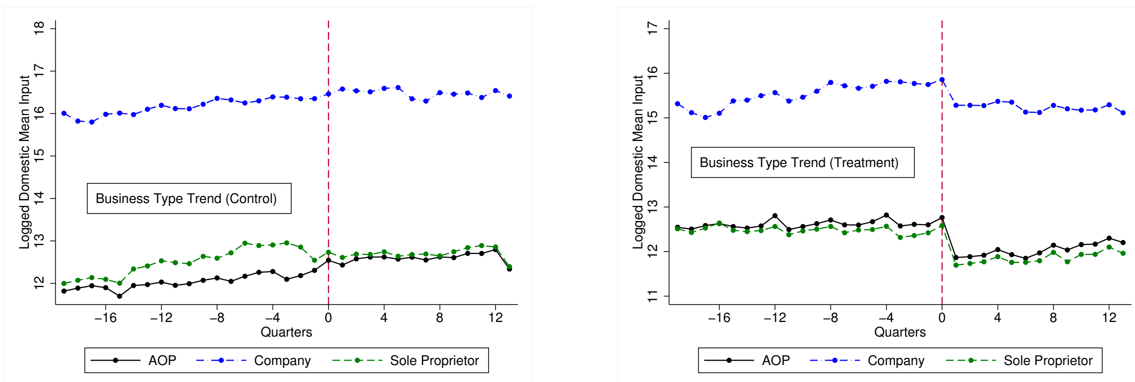
Figure A.1: Impact on Manufacturers vs. Non-Manufacturers by Business Type



Explanation: (Panel A & C) The graph plots the logged mean quarterly domestic input of manufacturers in control and treated groups based on their business type. The reform occurs at dashed vertical line (quarter April-June 2013) which is then used as a reference to show lead and lag quarter time periods. The input tax of each category drops after the reform. Decline in raw numbers is approximately 30 percent for the manufacturers in each category. (Panel B & D) The graph plots the logged mean quarterly domestic input of non-manufacturers in control and treated groups based on their business type. The reform occurs at dashed vertical line (quarter April-June 2013) which is then used as a reference to show lead and lag quarter time periods. The input tax of each category drops after the reform. Decline in raw numbers is approximately 70 log points or 50 percent for the companies and 90 log points or 60 percent for the partnerships and sole proprietorships.

Figure A.2: Heterogeneity by Business Types

(a) Control Group Response (b) Treatment Group Response



Explanation: The graph plots the logged mean quarterly domestic input of Companies, Partnerships and Individual Businesses in control and treated groups. The reform occurs at dashed vertical line (quarter April-June 2013) which is then used as a reference to show lead and lag quarter time periods. The input tax of each category drops after the reform. Decline in raw numbers is approximately 30 percent for the companies and 70 percent for individual and partnership businesses.

Table A.1: Descriptive Statistics for Sales, Purchases and Tax Credits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
Total Purchase	7.621	7.939	8.68	9.849	10.249	10.557	9.48	8.873	9.336
Taxable Purchase	6.437	7.313	8.015	9.373	9.785	10.067	9.106	8.432	8.797
Domestic Tax Credit	0.585	0.615	0.655	0.732	0.761	0.855	0.782	0.808	0.786
Import Tax Credit	0.226	0.248	0.292	0.374	0.347	0.349	0.371	0.432	0.426
Total Sale	9.304	9.526	10.547	13.13	13.528	13.591	13.402	11.77	12.242
Taxable Sale	8.2	8.591	9.63	10.437	10.644	10.307	9.508	8.631	8.954
Export sale	1.83	1.571	2.674	1.822	1.831	1.808	1.656	1.443	1.347
Observations	855,632	967,549	1,058,021	1,109,744	1,155,709	1,249,873	1,321,672	1,392,310	580,542

Notes: Table provides the financial yearly statistics of average purchase and sales for the eight complete years 2009-2016 and first five months of year 2017 in millions Pak Rupees (100 Pak Rupee = 1 US Dollar). The returns are filed on monthly basis except under very few special cases where the returns are required to be filed quarterly. Total purchase includes the exempt purchases as well as the taxable purchases. Taxable purchase is the total value of purchases including the one taxed at reduced or higher rate than the standard rate. Domestic tax credit is the input tax credit claimed against the purchases made locally and imported tax credit is the credit claimed against imports. Total sales include both exempt and taxable sales (including export sales which are zero rated).

Table A.2: Robustness to Monthly Time Periods (July 2008- September 2016)

	Domestic Input Tax (PKR in Millions)					
	(1)	(2)	(3)	(4)	(5)	(6)
DD (Post June 13 × Domestic Input Tax)	-1.15***					
	(0.3)					
DD (Post June 13 × Domestic Input Tax × Manufacturer)		-0.92***				
		(0.32)				
DD (Post June 13 × Domestic Input Tax × Non-Manufacturer)			-1.32***			
			(0.32)			
DD (Post June 13 ×Domestic Input Tax × Partnerships)				-1.44***		
				(0.3)		
DD (Post June 13 × Domestic Input Tax × Sole Proprietorships)					-1.38***	
					(0.31)	
DD (Post June 13 × Domestic Input Tax × Companies)						-1.89*
						(1.01)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes
Number of Groups	144,211	36,798	115,050	33,112	105,063	15,841
N	6,840,702	2,417,382	5,049,410	1,782,943	5,002,590	877,528

Notes: The table provides estimation of difference in difference coefficients for the specifications used in Table 1.2 to Table 1.7 for the complete period of July 2008 to September 2016 using the Monthly return data. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The coefficient estimates are approximately 1/3 of the coefficients in odd-numbered columns of Table A.3 because the time period is month instead of quarter. Standard Errors are clustered at firm level. The significance at 10%, 5% and 1% level is denoted by *, ** and *** respectively.

Table A.3: Robustness of the Results in Tables 2-7 to Full Period (July 2008- September 2016)

	Domestic Input Tax (PKR in Millions)											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Unbal	Bal	Unbal	Bal	Unbal	Bal	Unbal	Bal	UnBal	Bal	Unbal	Bal
DD (Post June 13 × Domestic Input Tax)	-3.37***	-3.44***										
	(0.86)	(0.9)										
DD (Post June 13 × Domestic Input Tax × Manufacturer)			-2.81***	-2.86***								
			(0.99)	(1.0)								
DD (Post June 13 × Domestic Input Tax × Non-Manufacturer)					-3.86***	-3.92***						
					(0.94)	(1.0)						
DD (Post June 13 × Domestic Input Tax × Partnerships)							-3.98***	-4.08***				
							(0.89)	(0.91)				
DD (Post June 13 × Domestic Input Tax × Sole Proprietorships)									-3.90***	-3.99***		
									(0.91)	(0.92)		
DD (Post June 13 × Domestic Input Tax × Companies)											-5.98**	-5.88**
											(2.98)	(3.01)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Groups	144,211	43,928	36,798	21,323	115,050	28,971	33,112	13,496	105,063	32,106	15,841	6,670
N	2,348,653	1,331,390	825,108	651,571	1,736,768	877,298	618,385	414,900	1,711,742	968,968	297,478	208,366

Notes: The table provides estimation of difference in difference coefficients for the specifications used in Table 1.2 to Table 1.7 for the complete period of July 2008 to September 2016. The Monthly return data is used to compute quarterly values, therefore N denotes the quarterly number of observations. The variable DD is defined as an interaction between the dummy for suppliers who were not claiming refund before July 2013 and the dummy which equals one for the period July 2013 onwards. The dependent variable is the input tax against domestic purchases and the regression controls for input tax against imports. The odd numbered columns show the results for the unbalanced panel which includes all the firms and the even numbered columns show the results for balanced panel of the firms used for Table 1.2 to Table 1.7. Standard Errors are clustered at firm level. The significance at 10%, 5% and 1% level is denoted by *, ** and *** respectively.

A.1.1 Legal Definitions & Institutional Background

A.1.1.1 Definitions

Many important terms such as distributor, companies etc. used in the paper are clearly defined under the law. All the definitions provided below are directly taken from the text of the Sales Tax Act, 1990. It should be noted here that VAT law was implemented in Pakistan without changing the name of the sales tax law which it replaced because of constitutional restrictions faced by the Federal Government in Pakistan.

I. **Association of Persons (AOP)** includes a firm, a Hindu undivided family, any artificial juridical person and anybody of persons formed under a foreign law, but does not include a company.

II. **Company** means – (a) a company as defined in the Companies Ordinance, 1984 (XL VII of 1984); (b) a body corporate formed by or under any law in force in Pakistan; (c) a modaraba; (d) a body incorporated by or under the law of a country outside Pakistan relating to incorporation of companies; (e) a trust, a co-operative society or a finance society or any other society established or constituted by or under any law for the time being in force; or (f) a foreign association, whether incorporated or not, which the Board has, by general or special order, declared to be a company for the purposes of the Income Tax Ordinance 2001 (XLIX of 2001)

III. **CREST** means the computerized program for analyzing and cross matching of sales tax returns, also referred to as COMPUTERISED RISK-BASED EVALUATION of SALES TAX

IV. **Distributor** means a person appointed by a manufacturer, importer or any other person for a specified area to purchase goods from him for further supply and includes a person who in addition to being a distributor is also engaged in supply of goods as a wholesaler or a retailer

V. **Supply Chain** means the series of transactions between buyers and sellers from the stage of first purchase or import to the stage of final supply

VI. **Wholesaler** includes a dealer and means any person who carries on, whether regularly or otherwise, the business of buying and selling goods by wholesale or of supplying or distributing goods, directly or indirectly, by wholesale for cash or deferred payment or for commission or other valuable consideration or stores such goods belonging to others as an agent for the purpose of sale

A.1.1.2 Invoice Summary Provision

Pakistan introduced federal VAT in 199o but with a very limited scope. In 1996, Pakistan expanded it and introduced standard credit invoice system VAT. The government intended to bring down excise and custom duties and expand tax base through a broad based consumption tax. Until 2001, the use of computers and software was minimal. The criminal elements exploited the zero rating against exports to defraud the government of billions of rupees through fake exports and invoices. Federal Board of Revenue (FBR) responded by launching STARR (Sales Tax Automated Refund Repository) in July 2002 which provided limited cross matching ability to the refund processing units. The criminal syndicates, however, continued to misuse, hack or dodge STARR. It also increased compliance cost for genuine firms significantly without curtailing the fraud substantially. A growing perception of inability of tax authorities to plug continued leakage put pressure on the government for more comprehensive measures. Consequently, FBR quickly moved to CREST in 2008. CREST enabled FBR to conduct more comprehensive risk analysis by scrutinizing transaction-level data through invoice summary filed as an annexure of monthly VAT return. FBR was able to capture the information that was previously unavailable, within risk analysis software automatically (Federal Board of Revenue 2008; Government of Pakistan 2008).

Pakistan's tax administration uses this transaction-level data to check fake input tax credit. The invoice summary provision in the tax law makes it mandatory for each VAT registered firm to file a monthly summary of purchase and sale invoices. The invoice summary, thus, gives digital synopsis of transactions. It includes registration number of each buyer and seller along with total number of invoices issued and the total tax involved in those invoices. This huge information is designed to limit different frauds including DMT fraud. The detailed format is in Appendix A.1.2 where the Annexure A, B, C, and D of the return show all the information captured in invoice summaries. Annexure A and C deal with purchases and sales respectively. Pakistani VAT regime requires compulsory electronic filing of VAT returns and its annexures. It implies that transaction-level data is available in electronic form for processing and counterchecking immediately with the filing of the return.

A.1.1.3 Missing Trader Fraud

The invoice summary provisions exist in most VAT regimes requiring the businesses to submit an electronic summary of sale and purchase invoices to substantiate their VAT return. The backward and forward linkage is designed to enable the tax authorities to comprehensively check the invoice trail in suspicious transactions. The non-deposit of input tax credit claimed on the basis of invoice issued by a non-existent seller can be denied retrospectively or through audit, making both the buyer and seller jointly and severally responsible for the deposit of tax.

DMT fraud operates in a chain. In Pakistani case, one firm issues invoices to the other and so on. Usually, the first supplier S_1, issues sales invoices of the desired goods to a buyer without actually supplying them. The buyer in these cases is a well-established business operating in formal sector, generally a manufacturer. The invoice issued by S_1 gives the buyer right to claim input tax credit although she actually purchased those goods from unregistered suppliers in the informal sector. In order to reduce her tax liability, the buyer now has legal claim of input tax against purchases, which never physically occurred. This can reduce tax liability of the buyer significantly. For example, a buyer who made purchases worth ten million PKR from the unregistered or informal sector can reduce her payable VAT by 1.5 million rupees (assuming a 15% tax rate). The self-enforcing mechanism of VAT demands that seller S_1 has a large amount of output tax which must be deposited in the treasury but to this end S_1 is backed by a chain of suppliers say S_2, S_3, S_4, S_5 etc. who can provide the fake input tax credit to reduce the actual tax payment by S_1 to zero or a negligible amount. One such network of suppliers who are criminally colluding with each other can deprive the exchequer to the tunes of billions of rupees each month.

These fake suppliers exploit the difficulty of audit and enforcement faced by the tax administration to get away with this fraud. The EU analogy is applicable here. In Pakistan, audit and enforcement jurisdictions are territorial and the auditors lack the authority and resources to conduct audit and verifications beyond their geographical limits. If the suppliers are carefully registered in different jurisdictions then these geographical limits work in a manner similar to the countries in EU but with far more ease of operation for the fraudsters. Clearly, if the suppliers S_1, S_2, ..., S_n are registered in different audit and enforcement jurisdictions, then practically there's very little an auditor can do. The investigation can be impeded further by two critical factors. First, the audit normally requires a period of year or more of activity and can take months or even years to complete and still more time is needed to get an enforceable order of recovery from the court. Second, once in the court, the courts are reluctant to buy the argument that based on a presumption some of the suppliers never existed at the time transaction took place. The government ends up giving refund or tax credit for the tax, which was never deposited in the treasury.

I elaborate it with an example. Suppose "M" is a manufacturer who buys recyclable paper and paperboard from large wholesalers operating in informal sector. It costs "M" ten million PKR to purchase this recyclable paper. M manufactures paper from it and sells it for PKR12.5 million. This firm is required to collect and remit a tax of 1,875,000 PKR (assuming a 15% tax rate) on this sale. If M can now get an invoice from S_1 for its purchase, then it reduces the tax liability by PKR 1,500,000. now collects full PKR 1,875,000 from its buyers but remits only 375,000 PKR. S_1 provides this fake invoice to M through a chain extending to S_2, S_3, S_4, S_5 and so on. The situation gets worse when M passes on some of this gain to the market through a reduced price. M starts capturing the market which leaves no other way for the competitors but to lower their cost by either engaging in similar fraud or changing its operations. Since the capital cost of changing operations is high and benefits are risky, the slippage to fraud is a more realistic and economically rational choice for the firm. This leads to an exponential growth where large segments of the industry get involved in these transactions. Virmani (1989) provides theoretical analysis of these types of problems in sales tax with reference to the evasion through mis-declaration of output but the intuition used by him can be extended to mis-declaration of inputs.

A.1.1.4 CREST and Reform in July 2013

CREST software analyzes and scrutinizes invoice summaries submitted by the buyers and suppliers. Then it goes back in the supply chain to identify any suspicious activity and points out invoice by invoice discrepancy. CREST is also linked with import and export data and cross verifies imports and export data submitted by the firms in their returns. Exhibit B-I provides a snapshot of the actual interface and output given by the CREST.

If an audit or further inquiry is necessary because either some invoices were not cleared by CREST or for any other reason, the amount cleared by CREST and approved by the refund processing division is sanctioned and the remaining amount is withheld pending further clarification. In short, refund claimant has to go through a month by month scrutiny which may often result in audit or inquiries. Through CREST system each invoice for the month is under scrutiny for refund claimant. This system is operational since financial year 2008 and the rules provide legal cover for the scrutiny of claims through CREST. Furthermore, the cases in which a firm is supplying goods locally as well as exporting them, CREST scrutinizes each and every invoice whether it pertains to a material used in export of goods or not. If CREST objects to a purchase, refund portion of the claim gets attenuated by the amount of

Figure A.3: CREST Output showing Supply Chain Discrepancies

Purchase Invoices

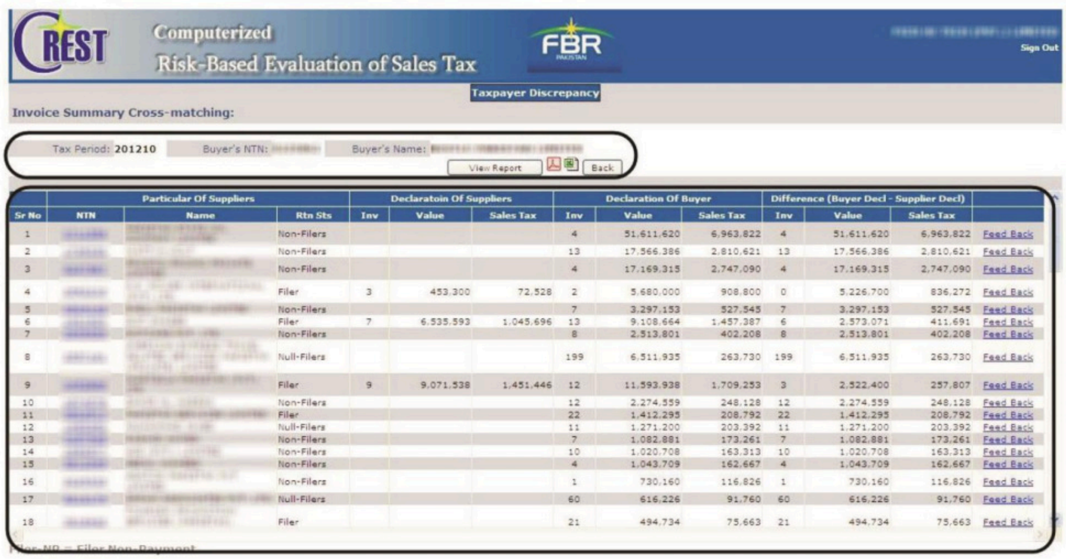
S.No.	Seller Reg. No.	Invoice No.	Invoice Date	Sales tax Amount	Status	Discrepancies	Verification Date
1.		676	02-01-14	4,143.00	Discrepancies	Taxable supply against refund filed not verified in supply chain	12-05-14
2.		5047	02-01-14	2,788.00	Discrepancies	Taxable supply against refund filed not verified in supply chain	12-05-14
3.		353	02-11-13	990.00	Valid		12-05-14
4.		628	04-12-13	30,888.00	Valid		12-05-14
5.		28019	04-01-14	2,141.00	Valid		12-05-14
6.		8	04-11-13	4,608.00	Discrepancies	Taxable supply against refund filed not verified in supply chain	12-05-14
7.		F41-0406	06-11-13	1,058.00	Valid		12-05-14
8.		26447	06-11-13	2,774.00	Valid		12-05-14
9.		F41-0546	07-01-14	1,498.00	Valid		12-05-14
10.		0869/01-14	07-01-14	1,475.00	Valid		12-05-14
11.		1365	07-01-14	28,377.00	Valid		12-05-14
12.		15174	07-01-14	6,720.00	Valid		12-05-14
13.		0100	07-01-14	3,390.00	Valid		12-05-14
14.		1367	07-01-14	31,417.00	Valid		12-05-14
15.		0063	07-11-13	3,865.00	Discrepancies	Taxable supply against refund filed not verified in supply chain	12-05-14
16.		90045341	09-01-14	54,400.00	Valid		12-05-14
17.		25	09-01-14	1,875.00	Valid		12-05-14
18.		699	09-01-14	1,900.00	Discrepancies	Taxable supply against refund filed not verified in supply chain	12-05-14
19.		4651	09-11-13	2,592.00	Valid		12-05-14
20.		28196	11-01-14	2,541.00	Valid		12-05-14
21.		0583/11-13	11-11-13	1,440.00	Valid		12-05-14
22.		9	12-11-13	4,760.00	Discrepancies	Taxable supply against refund filed not verified in supply chain	12-05-14

*** Discrepancy Raised by Processing Officer** **Page 1 of 4**

Explanation: The exhibit shows the output after processing of a refund claim through CREST along with invoice specific discrepancies raised against invoices with suspicious supply chain. This system was operational prior to the reform starting in early 2003 and completely rolled out countrywide by the financial year 2008.

that invoice even if the goods in question were not used in export. For example, a firm has total input for taxable purchases aggregating to PKR 1 million for a month but is only claiming a refund to the tune of PKR 0.5 million against exports. If CREST objects to PKR 0.1 million of input tax credit only, then the refundable amount takes the first hit and gets reduced to PKR 0.4 million.

Figure A.4: CREST Output Showing Other Discrepancies



	Particular Of Suppliers			Declaratoin Of Suppliers			Declaration Of Buyer			Difference (Buyer Decl - Supplier Decl)			
Sr No	NTN	Name	Rtn Sts	Inv	Value	Sales Tax	Inv	Value	Sales Tax	Inv	Value	Sales Tax	
1			Non-Filers				4	51,611,620	6,963,822	4	51,611,620	6,963,822	Feed Back
2			Non-Filers				13	17,566,386	2,810,621	13	17,566,386	2,810,621	Feed Back
3			Non-Filers				4	17,169,315	2,747,090	4	17,169,315	2,747,090	Feed Back
4			Filer	3	453,300	72,528	2	5,680,000	908,800	0	5,226,700	836,272	Feed Back
5			Non-Filers				7	3,297,153	527,545	7	3,297,153	527,545	Feed Back
6			Filer	7	6,535,593	1,045,696	13	9,108,664	1,457,387	6	2,573,071	411,691	Feed Back
7			Non-Filers				8	2,513,801	402,208	8	2,513,801	402,208	Feed Back
8			Null-Filers				199	6,511,935	263,730	199	6,511,935	263,730	Feed Back
9			Filer	9	9,071,538	1,451,446	12	11,593,938	1,709,253	3	2,522,400	257,807	Feed Back
10			Non-Filers				12	2,274,559	248,128	12	2,274,559	248,128	Feed Back
11			Filer				22	1,412,295	208,792	22	1,412,295	208,792	Feed Back
12			Null-Filers				11	1,271,200	203,392	11	1,271,200	203,392	Feed Back
13			Non-Filers				7	1,082,881	173,261	7	1,082,881	173,261	Feed Back
14			Non-Filers				10	1,020,708	163,313	10	1,020,708	163,313	Feed Back
15			Non-Filers				4	1,043,709	162,667	4	1,043,709	162,667	Feed Back
16			Non-Filers				1	730,160	116,826	1	730,160	116,826	Feed Back
17			Null-Filers				60	616,226	91,760	60	616,226	91,760	Feed Back
18			Filer				21	494,734	75,663	21	494,734	75,663	Feed Back

Filer-NP = Filer Non-Payment

Explanation: The exhibit shows the output for a supplier who is not a refund claimant by CREST along with invoice specific discrepancies raised against invoices with suspicious supply chain. This system was available to the tax authorities prior to the reform but didn't have any legal force in itself to deny input tax claim. After the reform in June 2013, the discrepancy raised by CREST against the domestic supplier as seen in the Exhibit automatically denied input tax claims against the difference amount.

A.1.2 Data Description

The data used in the paper is the administrative data of all the returns filed for the tax year 2009-2016 and first five months of the year 2017 constituting a total of 9.69 million returns. The data is anonymized so that it does not reflect actual registration numbers of firms to preserve confidentiality of specific firms. The returns are filed on monthly basis with the exception of a few industries or categories which file return on quarterly basis. Figure A.5 shows the return for the year 2016-17 (tax year 2017) on FBR's website and has several annexures which are also required to be filed with the return electronically. Figure A.5 only contains the annexures relevant for this paper. All returns are filed electronically, however, the period from 2009-2012 may have manual as well as electronic returns because the administration was transitioning towards electronic filing.

The data consists of 100 variables in total and few variables which directly relate to this paper are described in table A.4. Domestic input tax credit is the key variable. The penalty for not filing a return when no tax is due to be paid with the return is nominal throughout this period. The firms have the option to file a revised return to correct any error or misreporting in their return. A total of 3, 134 duplicate returns were filed and therefore dropped. The duplicate returns were also checked to ensure there's no change in the domestic input tax credit but no such case was found. Those who claimed refund in excess of 1 million PKR in total before the year of the reform were tagged as the control and those who obtained no or less than one million in total were tagged as treatment. The variable "Business Activity" includes all that apply, therefore, separate variables were generated to identify manufacturers and other business activities such as importer, wholesaler, and distributor (please refer to the definitions at Appendix A.1.1). The data for quarter Oct-Dec 2016 was not complete for the month of December and was therefore dropped from this analysis. Outliers in terms of domestic input tax credit were identified by setting a monthly threshold of PKR 2 billion to guard against data entry error. Only one such case was found and dropped. Missing values for the input tax credit for domestic as well as the import purchases were converted to 1 instead of zero for ease of calculations. It should be noted that a missing input tax value implies a zero claim.

Although the errors in data cannot be completely ruled out but the electronic filing on FBR's portal implies that the feeding errors that result in figure mismatches are eliminated. As one column of the return is calculated and links forward and backwards through in built software, data entry errors can be ignored. However, the firms can file a revised return, without prior approval voluntarily if that doesn't interfere with tax credits or payments such that tax liability remains the same or increases but in case the liability is to be revised downwards then a prior approval is required. The data does not show whether a duplicate return is revised or not but the duplicate returns are substantially less than 1% (3134 returns or 0.03%). For analysis purpose, I drop the duplicate returns for the same tax period but it is possible that revised return is dropped instead of the original one.

Table A.4: Data Variables and Description

Variable	Description
TAXPAYER_TYPE	Taxpayer Type (AOP/Company/Sole Proprietorship/FTN or Government Agencies)
BUSINESS_ACTIVITY	Business Activity (Manufacturer, distributor etc.), includes all that apply
ITEM_NAME	Name of the product sold, includes all that apply
CITY	City of registration
TAX_PERIOD	Monthly Tax Period in which return is filed
D_GPUPCH	Domestic Purchases from Registered Persons (excluding fixed assets) (Gross Value)
D_TPURCH	Domestic Purchases from Registered Persons (excluding fixed assets) (Taxable Value)
D_INPUT	Domestic input tax credit
DU_GPURCH	Domestic Purchases from Un-registered Persons (Gross Value)
I_GPURCH	Imports excluding fixed assets (includes value addition tax on commercial imports) (Gross Value)
I_TPURCH	Imports excluding fixed assets (includes value addition tax on commercial imports) (Taxable Value)
I_INPUT	Imported Input tax credit
FIX_GPURCH	Capital Goods / Fixed Assets (Domestic Purchases & Imports) (Gross Value)
FIX_TPURCH	Capital Goods / Fixed Assets (Domestic Purchases & Imports) (Taxable Value)
FIX_INPUT	Input Tax on account of Capital Goods / Fixed Assets (Domestic Purchases & Imports)
TOT_PURCH	Total Purchase (Gross Value)
TOT_TPURCH	Total Purchase (Taxable Value)
INPUT	Total Input tax credit for the month
STAX_CREDIT	Credit carried forward from previous tax period(s)
INADMIS_INPUT	Non creditable inputs (relating to exempt, non-taxed supplies of goods or services etc.)
D_GSALE	Total Goods or services supplied locally (Gross Value)
D_TSALE	Total Goods or services supplied locally (Taxable Value)
D_OUTPUT	Total Goods or services supplied locally (Sales Tax)
E_SALE1	Goods or Services exported (Gross Value)
TOT_SALE	Total Sales (Gross Value)
TOT_TSALE	Total Sales (Taxable Value)
G_OUTPUT	Output Tax
TURNOVER_TAX_BY_RETAILERS	Turnover Tax payable by retailers @ 2%
TO_OUTPUT	Retail Turnover - for the Quarter (Taxable Value)
TO_OUTPUT_TAX	Output Tax on Retail Turnover - for the Quarter
REFUND	Refund Claim (Provide Stock Statement as Annex-H)
TAX_PAYABLE	Total Tax Payable
TAX_PAID_NORMAL	Tax paid on normal/previous return (applicable in case of amended return)
BALANCE_TAX	Balance Tax Payable/ (Refundable)

Figure A.5: Sales Tax Return and Annexures

(a)

Government of Pakistan
Sales Tax & Federal Excise Return-cum-Payment Challan

"STR-7
[See rule 14(1)]

REGISTRY

NTN	CNIC (*in case of Individual*)	STRN (Sales Tax Registration No.)
-		

Name	Normal	Revised	Monthly	Quarterly	Tax Period (MMYY)

Section	No.	Description			Gross Value	Taxable Value	Sales Tax
SALES TAX CREDITS	1	Domestic Purchases (excluding fixed assets)	Annex-A		-	-	-
	2	Imports excluding fixed assets (includes 2% on commercial imports)	Annex-B		-	-	-
	3	Capital/ fixed assets to be credited at 1/12th of accumulated amount					-
	4	(-) Inadmissible input tax relating to exempt supplies/ non-taxed services etc.					-
	5	Input Tax for the month (1 + 2 + 3 - 4)					-
	6	(+) Previous month credit brought forward					-
	7	Accumulated Credit (5 + 6)					-
SALES TAX DEBITS	8	Supplies Made & Services Rendered	Annex-C		-	-	-
	9	Exports	Annex-D		-		
	10	Extra Tax charged under Chapter XIII of ST Sp. Procedure Rules '07	Annex-C				-
	11	Output Tax (8 + 10)					-
	12	Retail Turnover - for the Quarter			Turnover	-	-
	13	Electricity supplied to steel sector	KWH	-	x Rs. 6.00		-
	14	Re-rollable scrap sold by ship breakers	M Tons	-	x Rs. 4,848		-
	15	Re-meltable scrap sold by ship breakers	M Tons	-			
	16	Less: Sales Tax deducted by withholding agent @ 1/5th of tax invoiced					-
	17	Debit for the month (11 + 12 + 13 + 14 - 16)					-
	18	Sales Tax withheld by the return filer as withholding agent (STWH)					-
	19	Sales Tax Arrears including Principal, Def Surch. & penalty					-
PAYABLE/ REFUND	20	Whether excluded from Section 8B(1), under SRO 647(I)/2007				(Yes / No)	
	21	Admissible Credit - if 20 = Yes then 7; if 20 = No, then least of 7 or "90% of 11" or 17					-
	22	Payable ST - if 17 > 21 then (17 - 21 + 18 + 19); otherwise 18 + 19					-
	23	Excess Unadjusted Credit - if 20 = Yes and if 21 > 17 then (21 - 17); otherwise zero; if 20 = No then (7 - 21)					-
	24	Refund claim i.e. input consumed in zero-rated or excess of input tax as per rules					-
	25	Balance Credit to be carried forward - if 24 < 23, then (23 - 24); otherwizero zero					-
PAYMENTS	26	Federal Excise Duty (FED) Payable / (FED Drawback)				Annex-E	-
	27	Goods chargeable to Special Excise Duty (SED)				Annex-C	-
	28	(-) SED on inputs used in manufacturing of Goods supplied for domestic consumption					
	29	(-) SED paid on goods used in manufacturing of Goods exported (drawback)					
	30	Net SED Payable (27 - 28 - 29)					-
	31	SED Arrears					
	32	Net FED Payable If 26 + 30 > 0 then (26 + 30 + 31), else 31					-
	33	FED/ SED Drawback if 26 + 30 < 0 then -(26 + 30), else zero					-
	34	PDL - Petroleum Development Levy					-
	35	Total Taxes Payable (22 + 32+ 34)					-
	36	Tax paid on normal/ original return (applicable in case of revised return)					-
	37	Balance Tax Payable/ (Refundable) (35 - 36)					-
	38	Bank Account for payment of refund	A/C		Bank		Branch

DECLARATION

I,____________________________, holder of CNIC No. ________________

In my capacity as self/member or partner of association of persons/principal/ officer / trustee/ representative of named above, do solemnly declare that to the best of my knowledge and belief the information given in this return is correct and complete in accordance with the provisions of the Sales Tax Act, 1990, the Federal Excise Act, 2005, and rules as well as notifications issued thereunder.

*Date*__________ *Stamp*________________ *Signature*________________

HEAD OF ACCOUNTS

Head of Account	Amount	For Bank Use	CPR No.	Amount
B02341 - Sales Tax	-			-
B02366 - Sales Tax on services	-			-
B02367 - FED in VAT mode	-			-
B02485 - Federal Excise Duty	-			-
C03901 - PDL	-		Total Amount in Figures	-
TOTAL AMOUNT PAYABLE	-		Amount Received in words: ________ Bank Officer's Signatures, Date & Stamp	

Annex-A

SUMMARY OF DOMESTIC PURCHASES

NTN 99999999-9 | STRN xx-xx-xxxx-xxx-xx | Tax Period MM-YY

Name of Registered Person

S. Nr.	Particulars of Suppliers				No. of Invoices/ Debit/ Credit Notes	Value of Purchases Excluding S/Tax	Sales Tax	1% SED	Extra tax (paid under Chapter XIII of ST Sp. Procedures Rules, 2007)
	Name	NTN	STRN	CNIC					
Total (Net after incorporating the Debit/ Credit Notes, if any)						-	-	-	-

CATEGORY WISE SUMMARY:

			Value	Sales Tax
Taxable Goods				
@ 16% (excluding fixed assets)				-
@ 18.5% (excluding fixed assets)				-
@ 21% (excluding fixed assets)				-
Fixed Assets				
Third Schedule Goods				-
Taxable Services purchased (including provincial tax and FED in Sales Tax mode)				
@ 16%				-
@ 21%				-
Others (Pl. specify)				
Zero-rated				
DTRE				
Other local zero-rated				
Exempt domestic purchases/services				
Steel Sector: Tax paid at Rs. 6/KWH on electricity bill	**KWH**			
CNG Dealers: Natural gas purchased	**Tax paid at 25%**			

Notes:

1) Supplier-wise summary should be provided for all taxable (excluding zero-rated) purchases made from registered

2) Supplier-wise summary is not required to be submitted by retailers and CNG dealers and in respect of those purchases on

3) 'Others' category also covers purchases made from unregistered persons and should also include purchases for which no separate column is provided.

(b)

Annex-B

SUMMARY OF IMPORTS

NTN 99999999-9 | STRN xx-xx-xxxx-xxx-xx | Tax Period MM-YY

Name of Registered Person

S. Nr.	Particulars of GD Imports (Machine No.)				Sales Tax Rate	Import Type	Value for Sales Tax	Sales Tax paid at import stage	2% Sales Tax on commercial imports	1% SED
	Collectorate	GD Type	GD No.	GD Date						
Total							-	-	-	-

GD Types	Collectorates	Import Types
EB	KAPR	General
HC	KOIL	Commercial
ST Rates	KAFU	Fixed Assets
0%	LDRY	Edible Oil
16%	PDRY	Ship for breaking
18.5%	MDRY	
21%	QDRY	
Exempt	(Also see instructions)	

Rate/ Type-wise summary		Value	Sales Tax
Imports @ 16% (excl. fixed assets)			-
Imports @ 18.5% (excl. fixed assets)			-
Imports @ 21% (excl. fixed assets)			-
Zero Rated			
Imports at other rates			-
Fixed Assets			-
Exempt Imports			
Ship for breaking	LDT		

Note: If there are if items of different rates and types are imported on a single GD, separate line for each type/ rate may be entered in GD-wise summary

(c)

Annex-C

SUMMARY OF DOMESTIC SALES

NTN 99999999-9 STRN xx-xx-xxxx-xxx-xx Tax Period MM-YY

Name of Registered Person

S. Nr.	Particulars of buyers				No. of Invoices/ Debit/ Credit Notes	Value of Supplies Excluding S/Tax	Sales Tax	1% SED	Extra tax (charged under Chapter XIII of ST Sp. Procedures Rules, 2007)
	Name	NTN	STRN	CNIC					
	All supplies to unregistered persons								
Total (Net after incorporating the Debit/ Credit Notes, if any)						-	-	-	-

CATEGORY WISE SUMMARY:

	Value	Sales Tax
Taxable Goods & services		
@ 16%		-
@ 18.5%		-
@ 21%		-
@ 25% (Natural Gas supplied to CNG dealers)		-
Third Schedule Goods		-
Taxable Services rendered (including provincial tax and FED in Sales Tax mode)		
@ 16%		-
@ 21%		-
Others (Pl. specify)		-
Zero-rated		
DTRE		
Other local zero-rated		
Exempt supplies/services		
Invoices issued under special procedures — **Tax invoiced**		

Notes:

1) The buyer-wise summary should be provided for all taxable sales (excluding zero-rated) to registered persons.
2) The buyer-wise summary is not required to be submitted by retailers and CNG dealers.
3) 'Others' category covers supplies for which no separate column is provided.
4) 'Invoices issued under special procedures' reflect sales tax for which sales tax liability is discharged under special procedures and tax on invoice does not form part of output tax.

(d)

Annex-D

SUMMARY OF EXPORTS

NTN 99999999-9 STRN xx-xx-xxxx-xxx-xx Tax Period MM-YY

Name of Registered Person

S. Nr.	Particulars of GD Exports (Machine No.)				Value of Exports in Pak Rupees
	Collectorate	GD Type	GD No.	GD Date	
Total					-

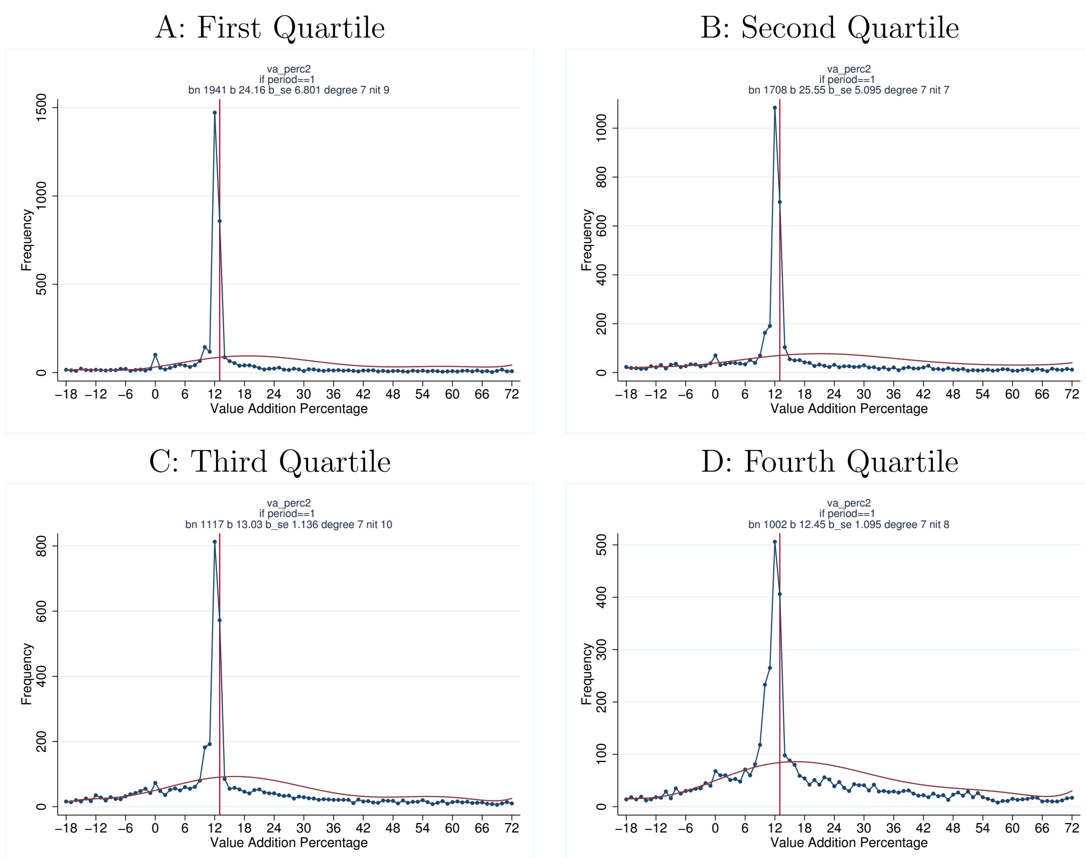
(e)

Appendix B

Appendix to Chapter 2

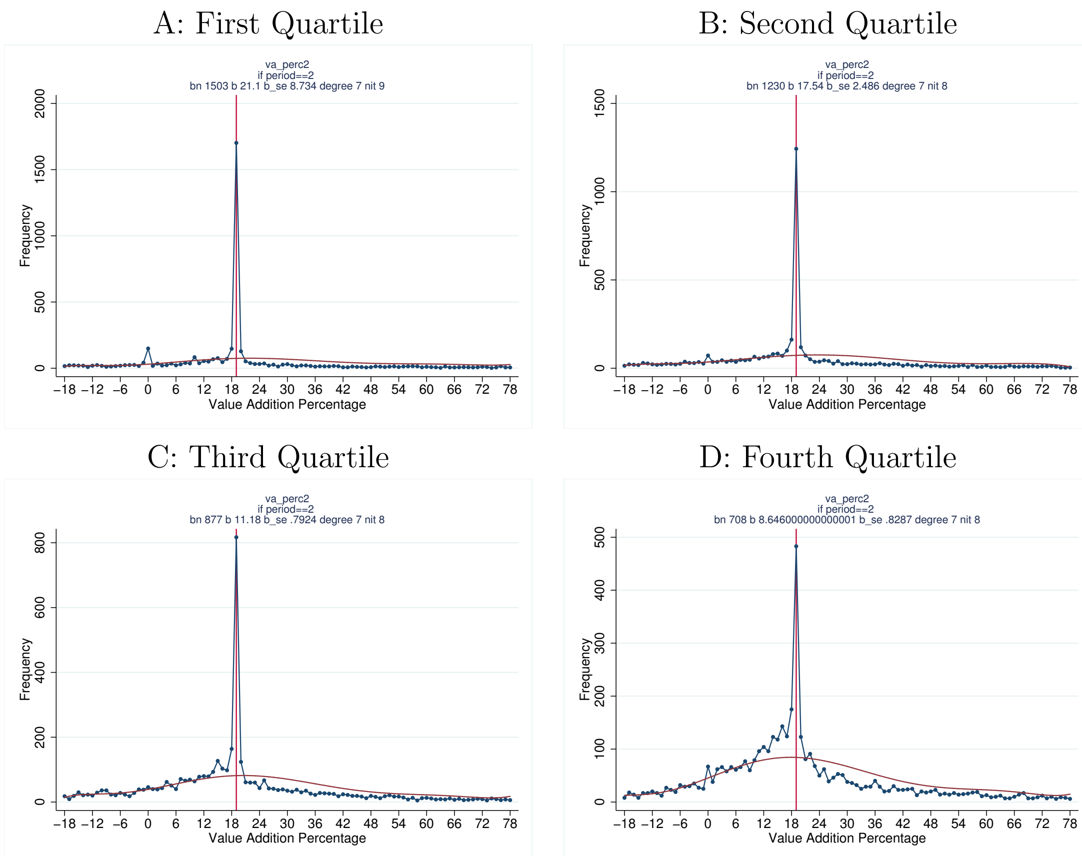
B.1 Additional Tables and Figures

Figure B.1.1: Quartile Wise Bunching Estimates for First Period



Explanation: The graph plots number of firms along with its counterfactual distribution, showing bunching for different quartiles during first period. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 1 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 5% would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. The tails on left and right of bunching window are relatively thick for fourth quartile suggesting that these firms possibly face more adjustment costs.

Figure B.1.2: Quartile Wise Bunching Estimates for Second Period



Explanation: The graph plots number of firms along with its counterfactual distribution, showing bunching for different quartiles during second period. Horizontal axes show value addition percentage which is calculated using gross sales and imports in a year for each importing firm. Each bin on this axis has a width of 1 percent such that, for example, any firm showing value addition equal to or greater than 4% but less than 5% would be counted in that bin. Vertical axes show number of firms in the bins. Red vertical lines show MVA threshold for a particular year. The tails on left and right of bunching window are relatively thick for fourth quartile suggesting that these firms possibly face more adjustment costs.

B.1.1 Data Description

I use administrative data of all the returns filed for the tax year 2009-2016 and first five months of the year 2017 constituting a total of 9.69 million returns. Normal returns are filed on monthly basis with the exception of a few industries or categories which file return on quarterly basis or may file special or revised returns. The data consists of 100 variables in total and few variables which directly relate to this chapter are described in table B.1.1. Import value and domestic sale are key variables. The return has a separate columns for input tax paid as MVA tax at time of import. Taxable sales value may include some taxable purchase and sales at domestic level but they are not bifurcated for the analysis in this essay.

Although the errors in data cannot be completely ruled out but the electronic filing on FBR's portal implies that the feeding errors that result in figure mismatches are eliminated. As one column of the return is calculated and links forward and backwards through in built software, data entry errors can be ignored. However, the firms can file a revised return, without prior approval voluntarily if that doesn't interfere with tax credits or payments such that tax liability remains the same or increases but in case the liability is to be revised downwards then a prior approval is required. The data does not show whether a duplicate return is revised or not but the duplicate returns are substantially less than 1% (3134 returns or 0.03%). For analysis purpose, I drop the duplicate returns for the same tax period but it is possible that revised return is dropped instead of the original one.

Table B.1.1: Data Variables and Description

Variable	**Description**
TAXPAYER_TYPE	Taxpayer Type (AOP/Company/Sole Proprietorship/FTN or Government Agencies)
BUSINESS_ACTIVITY	Business Activity (Manufacturer, distributor etc.), includes all that apply
ITEM_NAME	Name of the product sold, includes all that apply
CITY	City of registration
TAX_PERIOD	Monthly Tax Period in which return is filed
D_GPUPCH	Domestic Purchases from Registered Persons (excluding fixed assets) (Gross Value)
D_TPURCH	Domestic Purchases from Registered Persons (excluding fixed assets) (Taxable Value)
D_INPUT	Domestic input tax credit
DU_GPURCH	Domestic Purchases from Un-registered Persons (Gross Value)
I_GPURCH	Imports excluding fixed assets (includes value addition tax on commercial imports) (Gross Value)
I_TPURCH	Imports excluding fixed assets (includes value addition tax on commercial imports) (Taxable Value)
I_INPUT	Imported Input tax credit
FIX_GPURCH	Capital Goods / Fixed Assets (Domestic Purchases & Imports) (Gross Value)
FIX_TPURCH	Capital Goods / Fixed Assets (Domestic Purchases & Imports) (Taxable Value)
FIX_INPUT	Input Tax on account of Capital Goods / Fixed Assets (Domestic Purchases & Imports)
TOT_PURCH	Total Purchase (Gross Value)
TOT_TPURCH	Total Purchase (Taxable Value)
INPUT	Total Input tax credit for the month
STAX_CREDIT	Credit carried forward from previous tax period(s)
INADMIS_INPUT	Non creditable inputs (relating to exempt, non-taxed supplies of goods or services etc.)
D_GSALE	Total Goods or services supplied locally (Gross Value)
D_TSALE	Total Goods or services supplied locally (Taxable Value)
D_OUTPUT	Total Goods or services supplied locally (Sales Tax)
E_SALE1	Goods or Services exported (Gross Value)
TOT_SALE	Total Sales (Gross Value)
TOT_TSALE	Total Sales (Taxable Value)
G_OUTPUT	Output Tax
TURNOVER_TAX_BY_RETAILERS	Turnover Tax payable by retailers @ 2%
TO_OUTPUT	Retail Turnover - for the Quarter (Taxable Value)
TO_OUTPUT_TAX	Output Tax on Retail Turnover - for the Quarter
REFUND	Refund Claim (Provide Stock Statement as Annex-H)
TAX_PAYABLE	Total Tax Payable
TAX_PAID_NORMAL	Tax paid on normal/previous return (applicable in case of amended return)
BALANCE_TAX	Balance Tax Payable/ (Refundable)

Appendix C

Appendix to Chapter 3

C.1 Definition of Variables

(*i*) **Sales.** The value of all goods and services supplied by the firm in the given tax period (month) including exports.

(*ii*) **Purchases.** The value of all taxable intermediates acquired by the firm in the given tax period (month).

(*iii*) **Output Tax.** The value of VAT charged on sales made by the firm in the given tax period (month). It equals $\tau.(\hat{s}_{it} - \hat{s}_{E,it})$, where τ is the applicable VAT rate and $(\hat{s}_{it} - \hat{s}_{E,it})$ is the value of non-export sales reported by firm i in period t. Because exports are zero-rated, they do not appear in the output tax.

(*iv*) **Input Tax.** The value of VAT credit claimed on intermediates acquired by the firm in the given tax period (month). It equals $\tau.\hat{c}_{it}$, where τ is the applicable VAT rate and $\hat{c}_{it}$ is the value of purchases of intermediates claimed by firm i in period t.

(*v*) **Tax Payable.** The VAT payable by the firm in the given tax period (month). By definition, it equals the output tax minus the input tax.

(*vi*) **Tax Paid** The VAT actually paid by the firm in the given tax period (month). It may differ from Tax Payable if the firm has any carry-forward from previous months.

(*vii*) **Exports.** The value of all goods and services exported by the firm in the given tax period (month).

(*viii*) **Imports.** The value of all goods and services imported by the firm in the given tax period (month).

(*ix*) **Refund.** The amount of refund claimed by the firm in the given tax period (month). The refund arises when the firm's input tax exceeds its output tax. In this case, the firm has the option to carry forward the balance amount or seek its refund. Because exports are zero-rated, firms the majority of whose output is exported are likely to claim refund every tax period.

(*x*) **Carry Forward.** The amount of carry forward claimed by a firm. The carry forward arises when the firm's input tax exceeds its output tax and it does not opt to seek the refund of the balance amount.

(*xi*) **Manufacturer.** A firm whose principal business activity is the manufacture of goods. Manufacturing is the process whereby a firm converts inputs into a distinct article capable of being put to use differently than inputs and includes any process incidental or ancillary to it.

(*xii*) **Importer.** A firm whose principal business activity is the import of goods for sale in the local market without carrying out any manufacturing process on them.

(*xiii*) **Exporter.** A firm whose principal business activity is the export of goods. These firms may supply in the local market, but a majority of their output is exported out of country.

(*xiv*) **Distributor.** Distributor means a person appointed by a manufacturer, importer or any other person for a specified area to purchase goods from him for further supply and includes a person who in addition to being a distributor is also engaged in supply of goods as a wholesaler or a retailer.

(*xv*) **Wholesaler.** Wholesaler' includes a dealer and means any person who carries on, whether regularly or otherwise, the business of buying and selling goods by wholesale or of supplying or distributing goods, directly or indirectly, by wholesale for cash or deferred payment or for commission or other valuable consideration or stores such goods belonging to others as an agent for the purpose of sale; and includes a person supplying taxable goods to a person who deducts income tax at source under the Income Tax Ordinance, 2001.

(*xvi*) **Retailer.** A person, supplying goods to general public for the purpose of consumption.

($xvii$) **Industry.** The Pakistani tax administration uses 4-digit Harmonized Commodity Description and Coding System (HS code) to classify firms into industry. The code, used by customs administrations throughout the world, divides all goods and services into 99 chapters (the first two digits in the code) and 21 sections. The sections broadly correspond to major industries in the country. I take the section a firm falls in as its industry.

($xviii$) **Major City** The dummy variable takes the value 1 if the firm's head office is in one of the three major cities of Pakistan—Karachi, Lahore, and Islamabad.

(xix) **LTU** The dummy variable takes the value 1 if the firm is administered by on of the four Large Taxpayer Centers in the country located in Karachi, Lahore, and Islamabad.

C.2 Marginal Treatment Effects

In this section, we describe how we estimate the $MTE(p)$ curves shown in Figures 3.9 and 3.10. Because we have access to a binary instrument only, full nonparametric identification (see James J. Heckman and E. Vytlacil 2005; James J. Heckman and Edward J. Vytlacil 2007) is not feasible in our setup, and instead we identify MTEs under a functional structure following the approach developed in Kowalski (2016) and Brinch, Mogstad, and Wiswall (2017).

As in the paper, Z here denotes the instrument (random assignment) and D the treatment (actual audit). Following the standard terminology in this literature, we refer to $p \equiv P(D = 1|Z)$ as the potential fraction treated. For any outcome Y, The $MTE(p)$ is defined as

$$MTE(p) \equiv \mathbb{E}(Y_T - Y_U | U_D = p)$$

where Y_T represents the potential outcome in the audited state ($D = 1$) and Y_U the potential outcome in the unaudited state ($D = 0$). The unobserved cost and benefit of audit are represented by U_D and p. The MTE therefore captures the treatment effect on a unit marginal to selecting into treatment. Using the above definition, it can be written as the difference between the marginal treated outcome (MTO) and the marginal untreated outcome (MUO)

$$\begin{aligned} MTO(p) \equiv &\ \mathbb{E}(Y_T | U_D = p) \\ MUO(p) \equiv &\ \mathbb{E}(Y_U | U_D = p) \end{aligned}$$

These curves are defined for every value of $p(Z)$ but given our binary instrument only two values of p are observed: the baseline treatment probability $p_B \equiv P(D=1|Z=0)$ and the intervention treatment probability $p_I \equiv P(D=1|Z=1)$. We therefore assume that both these curves are linear. The $MTO(p)$ is identified at two points

$$\begin{aligned} BTTO = & \quad \mathbb{E}(Y|X=x, D=1, Z=0) \\ LATO = & \quad \tfrac{1}{p_I - p_B}\left[p_I ITTO - p_B BTTO\right], \end{aligned}$$

where $ITTO = \mathbb{E}(Y|X=x, D=1, Z=1)$. We use the linearity assumption to extrapolate between these two points. Similarly, the $MUO(p)$ is identified at

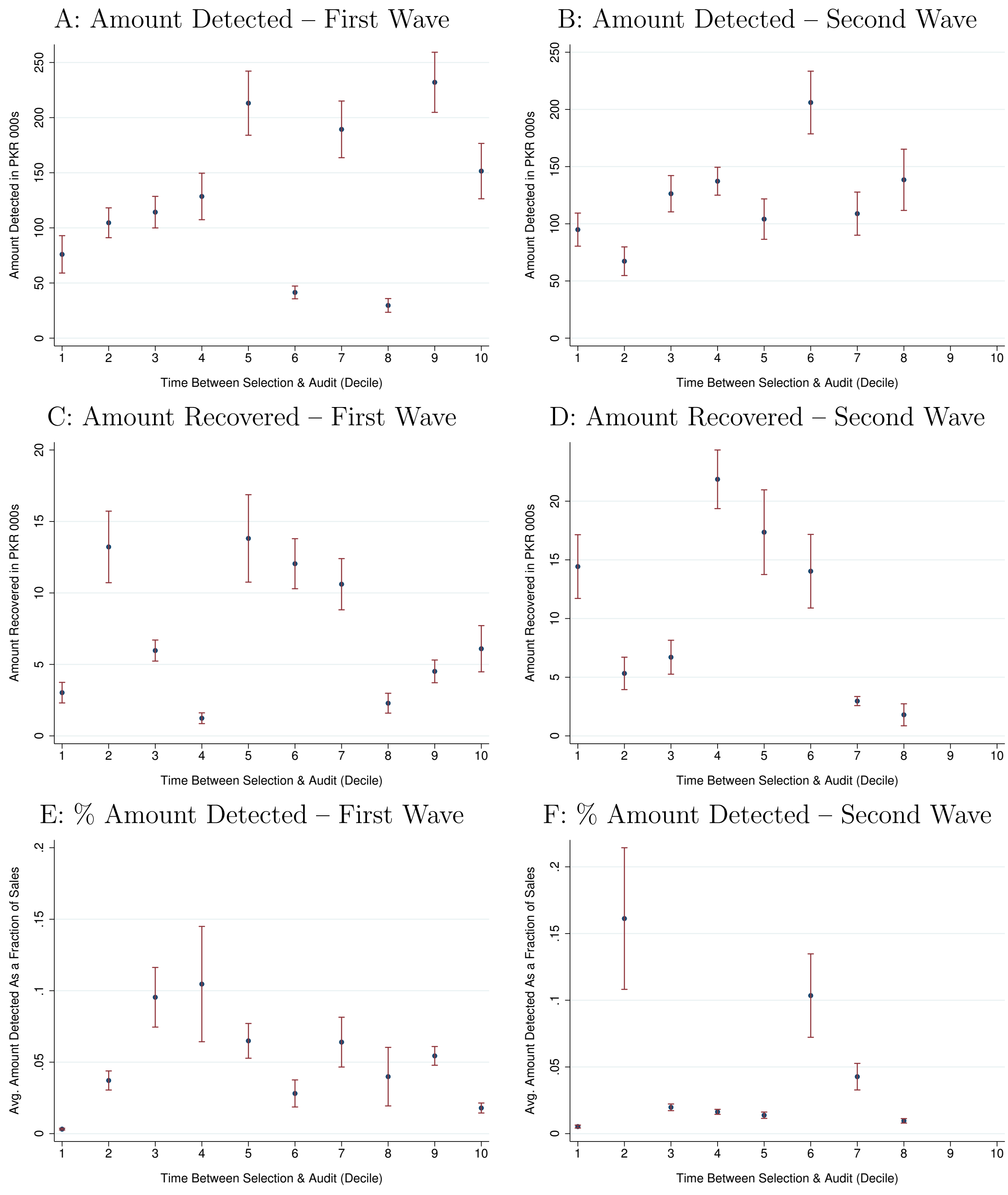
$$\begin{aligned} IUUO = & \quad \mathbb{E}(Y|X=x, D=0, Z=1) \\ LAUO = & \quad \tfrac{1}{p_I - p_B}\left[(1-p_B)BUUO - (1-p_I)IUUO\right] \end{aligned}$$

where $BUUO = \mathbb{E}(Y|X=x, D=0, Z=0)$.[1]

To plot the $MTO(p)$ curve, we regress the outcome variable on a full set of firm and period fixed fixed effects and an interaction term of the audit (D) and post dummies, restricting the sample to firms randomly selected for audit ($Z=1$). The regression gives us estimates of ITTO and IUUO. Running a similar regression on a sample of firms not drawn in the random ballot ($Z=0$) delivers the estimates of BTTO and BUUO. We then find LATO and LAUO using the definitions above. The $MTO(p)$ curve is identified at two points $(BTTO, \frac{p_B}{2})$ and $(LATO, \frac{p_B+p_I}{2})$. We extrapolate between the two using the linearity assumption. Similarly, $MUO(p)$ is identified at $(LAUO, \frac{p_B+p_I}{2})$ and $(IUUO, \frac{p_I+1}{2})$, and we extrapolate using linearity. The $MTO(p)$ curve is the difference between the two. We draw these curves for four outcomes and two audit waves separately. Since in our setting all these curves sit above each other, we lift both $MTO(p)$ and $MUO(p)$ up by adding the constant from the corresponding regression to distinguish them from the primary object of our interest $MTE(p)$.

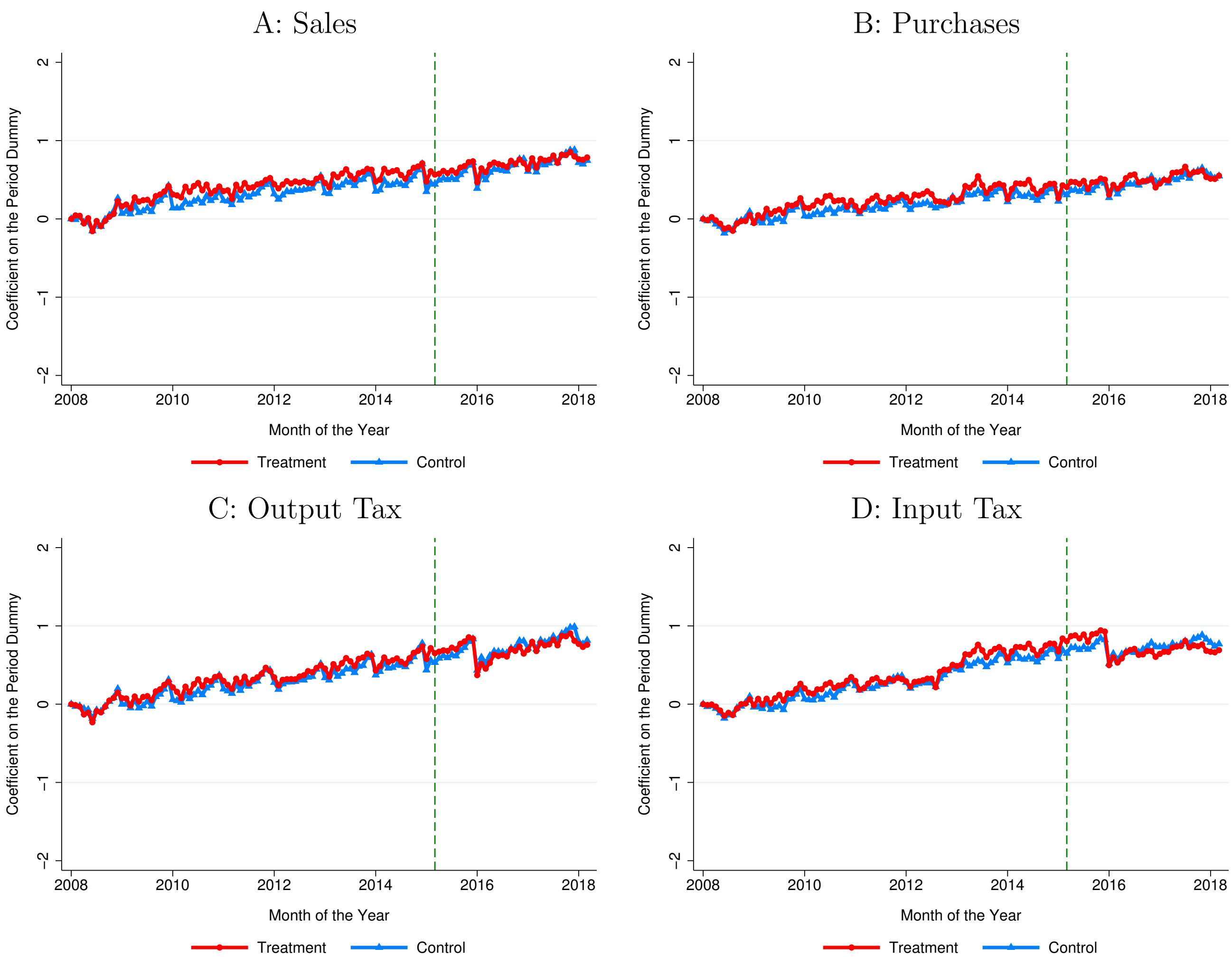
[1]In all these definitions, O stands for outcomes, T for treated, U for untreated, B for baseline, I for intervention, and LA for local average. Please see **kowalski2016doing** for detail of these terms.

Figure C.2.1: Amount Detected by Timing of Audit



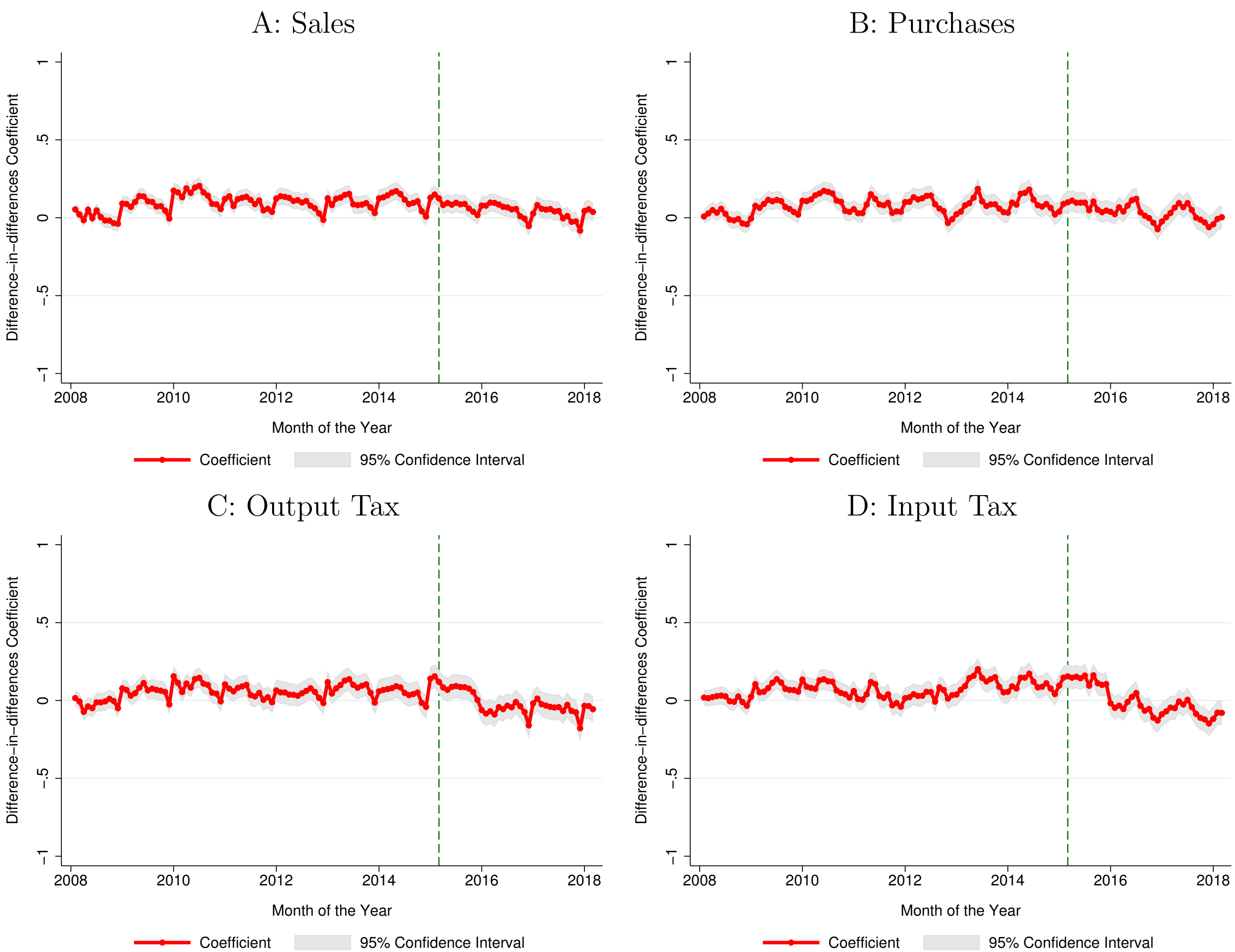
Notes: The figure examines if the order in which audits were taken up is correlated with audit outcomes, exploring thereby if audits were systematically targeted toward specific firms. We divide the time between assignment and initiation of audit into ten deciles and then plot the average audit outcome and the 95% confidence interval around it for each decile. The top panels look at the average amount detected by audit in PKR thousands, the middle panels at the average amount recovered in PKR thousands, and the bottom panels at the average amount detected as a ratio of annual baseline turnover of the firm. To take care of outliers, we drop observations where the amount detected is more than the 99th percentile of the distribution. This affects the top and bottom panels only. The LHS panels plot outcomes for the first randomized ballot and the RHS for the second.

Figure C.2.2: Intention to Treat Effects of Third Audit Wave



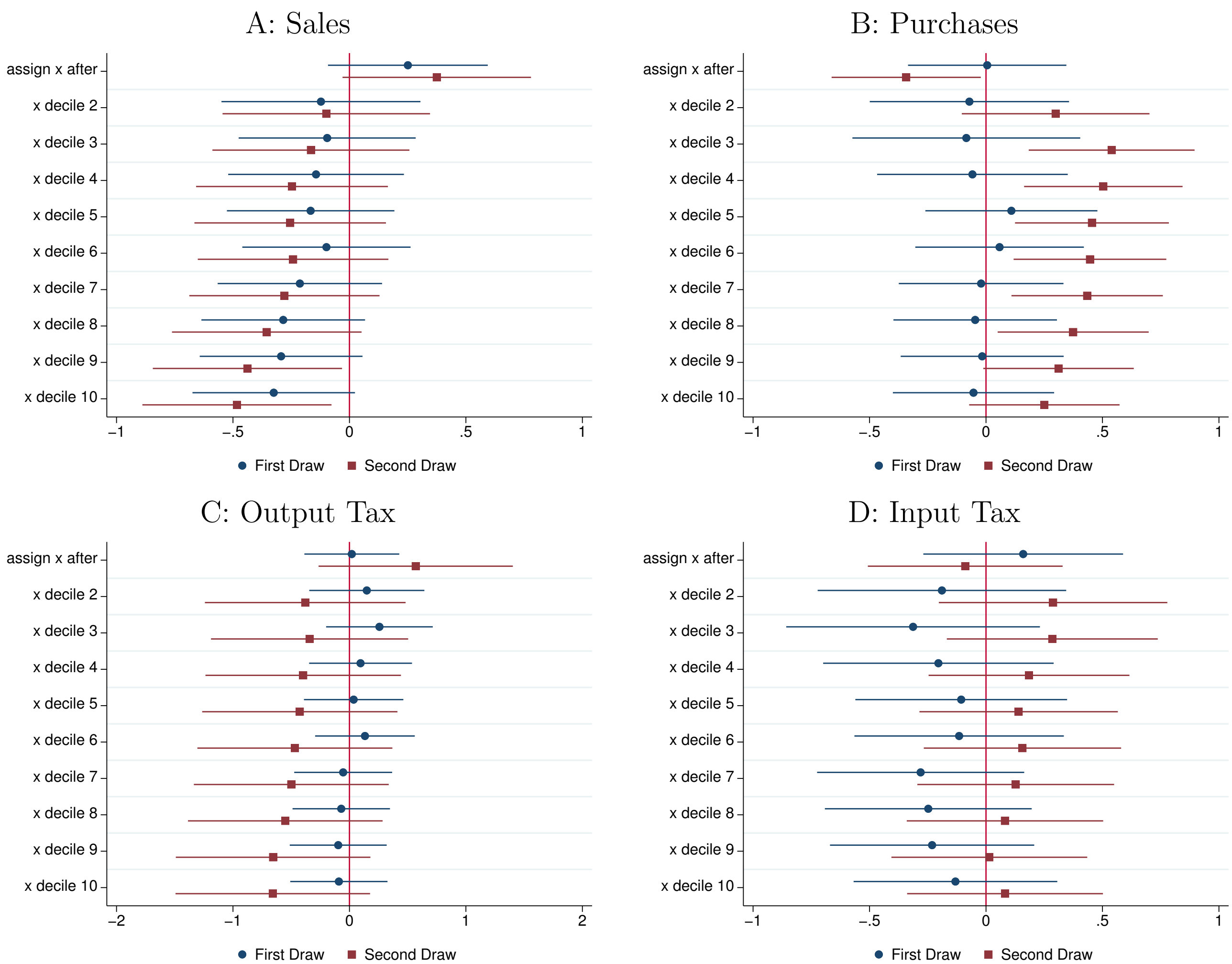
Notes: The figure explores the impacts of audit on future firm behavior. We compare the evolution of four VAT outcomes across the treatment and control groups. The treatment groups consists of firms whose audit was assigned through the first random ballot held on September 14, 2015. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments, firms already under audit, and firms subject to fixed and withholding tax regimes. We do not identify the last type of firms and therefore are unable to exclude them from the eligible sample. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm and month fixed effects, dropping the dummy for July 2008. We then plot the coefficients on the time dummies of these regressions. The sample includes all tax periods from July 2008 to June 2018. The regressions are run separately for the two groups of firms. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines demarcate the date the random computer ballot was held on.

Figure C.2.3: Intention to Treat Effects of Third Audit Wave



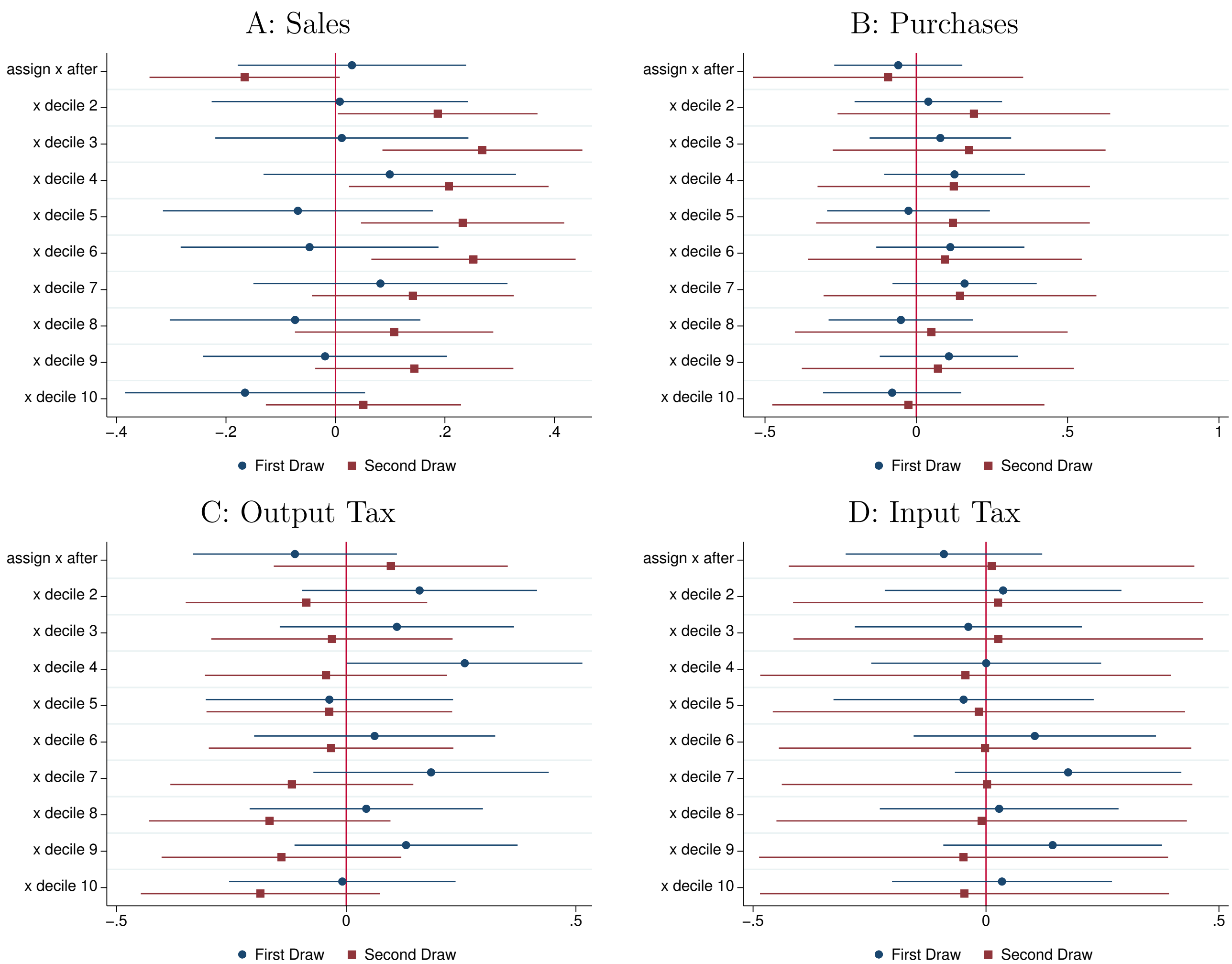
Notes: The figure shows the difference-in-differences version of the plots in Figure C.2.2. To construct these charts, we regress the log of the outcome variable shown in the title of each panel on the full set of firm, month, and month×treat dummies, dropping the dummies for July 2008. We then plot the coefficients on the month×treat dummies from these regressions. The gray surface plot shows the 95% confidence interval around the coefficient. The treatment groups consists of firms whose audit was assigned through the first random ballot held on September 14, 2015. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments, firms already under audit, and firms subject to fixed and withholding tax regimes. We do not identify the last type of firms and therefore are unable to exclude them from the eligible sample. We cluster standard errors at the firm level. Year t on the horizontal axis indicates July of the corresponding year. Vertical dashed lines demarcate the date the random computer ballot was held on.

Figure C.2.4: Heterogeneity in Response by Firm Size



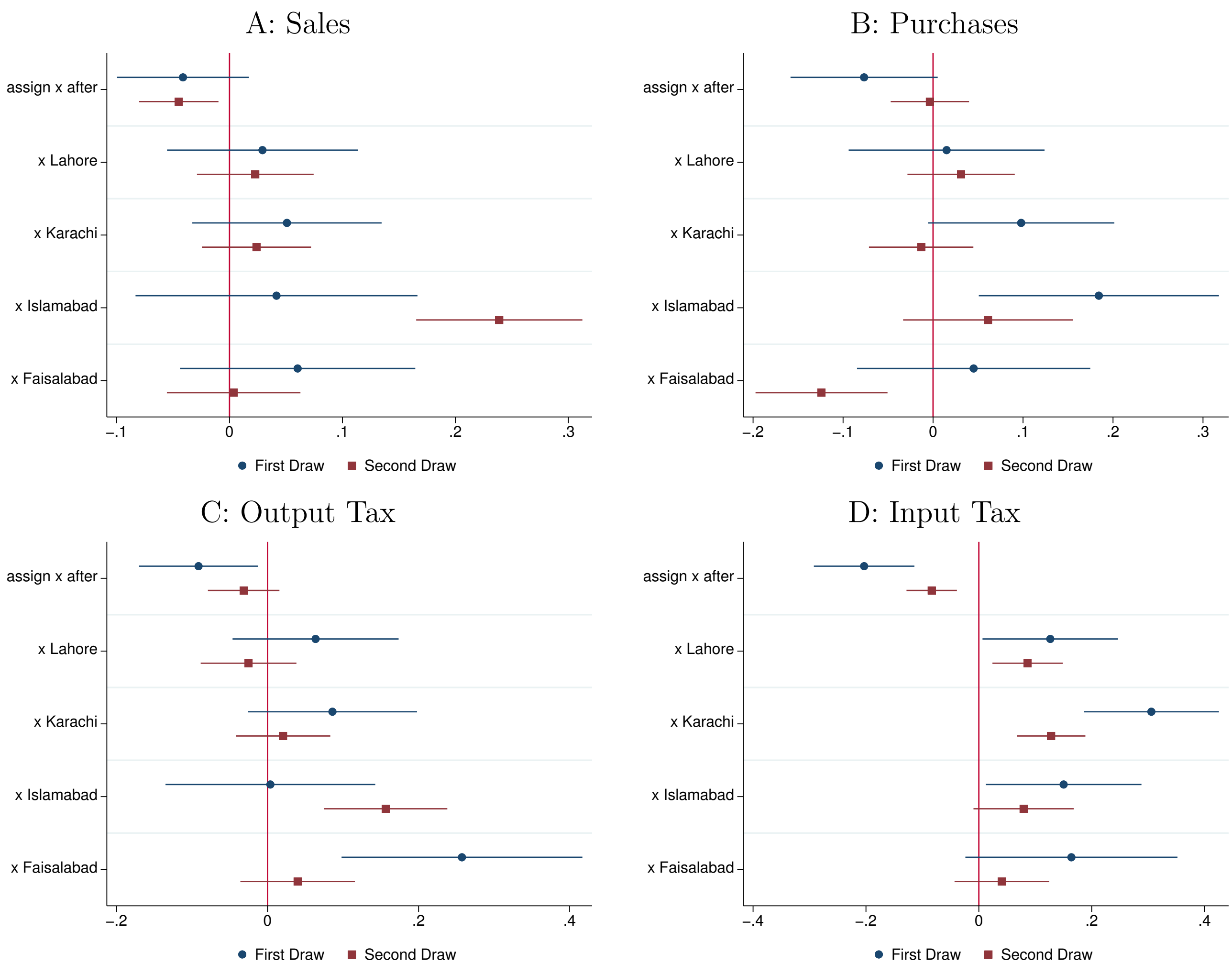
Notes: The figure explores heterogeneity in the audit effect. We divide firms into ten deciles based on their annual turnover in the baseline year. We then estimate a triple-difference version of model (3.10). The model includes interactions of the firm decile dummy with the $assign \times after_{it}$ dummy. The $assign_i$ dummies takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. We drop the triple-interaction term involving the first decile. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.5: Heterogeneity in Response by Firm Age



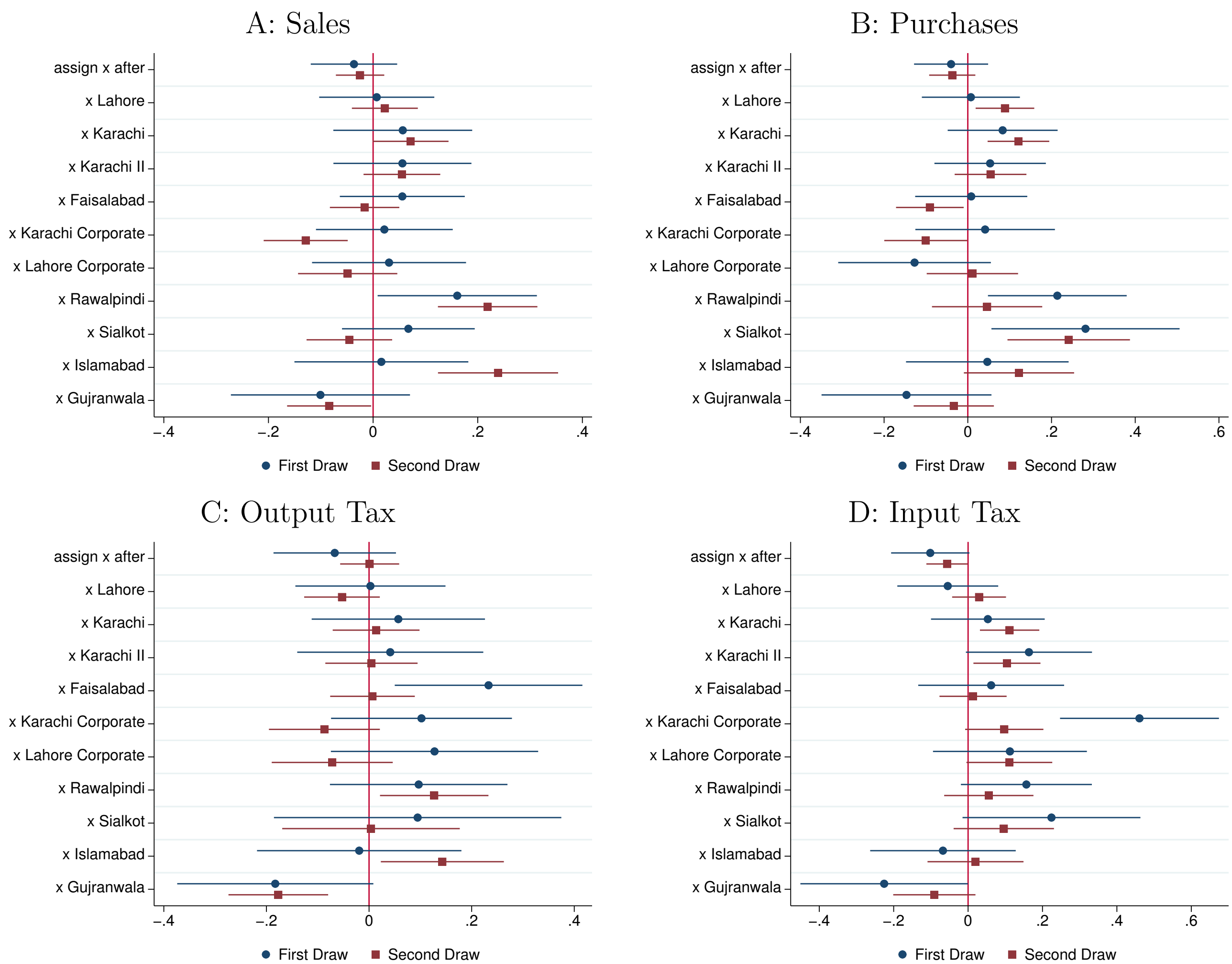
Notes: The figure explores heterogeneity in the audit effect. We divide firms into ten deciles based on their age, defining age as the number of days between July 1, 2013 and the date of registration of the firm. We then estimate a triple-difference version of model (3.10). The model includes interactions of the firm decile dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. We drop the triple-interaction term involving the first decile. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.6: Heterogeneity in Response by Location



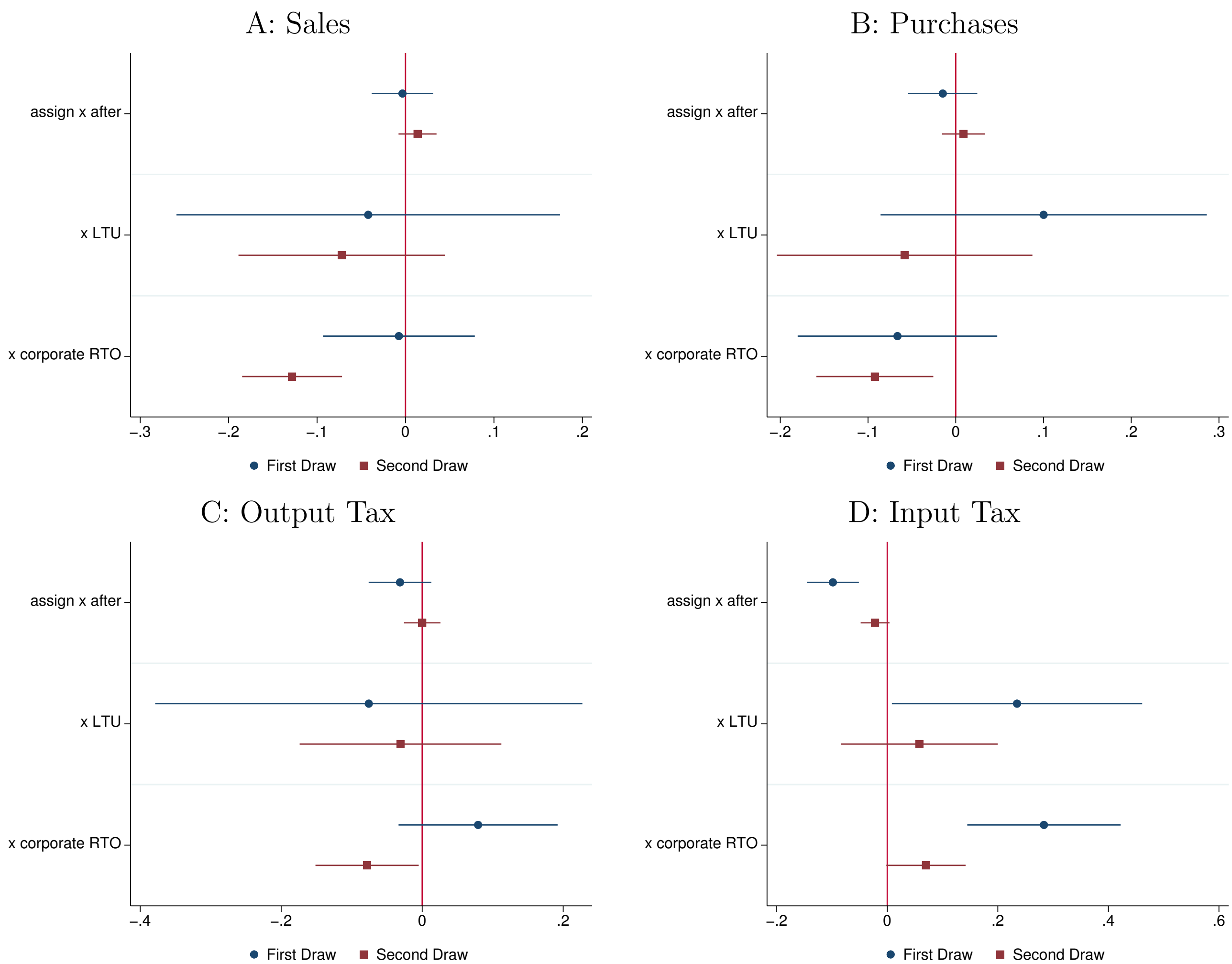
Notes: The figure explores heterogeneity in the audit effect. We divide firms into five groups depending upon the city their head office is located in. Firms not located in the four major cities of the country— Lahore, Karachi, Islamabad, and Faisalabad— are included in the baseline category. We then estimate a triple-difference version of model (3.10). The model includes interactions of the firm location dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.7: Heterogeneity in Response by Tax Office



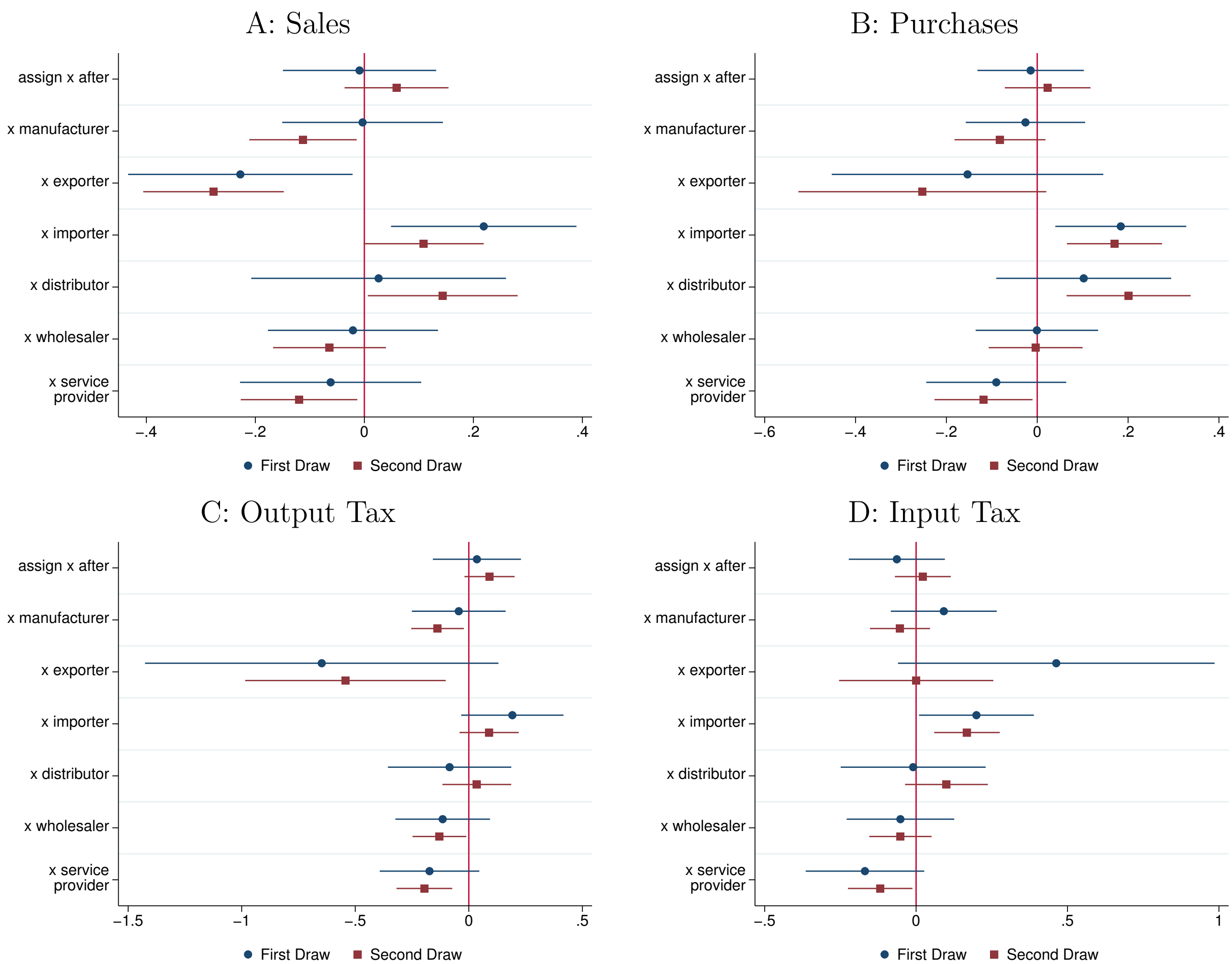
Notes: The figure explores heterogeneity in the audit effect. We divide firms into eleven groups based on the local tax office they are subject to. Firms not in the ten major tax offices are included in the baseline category. We then estimate a triple-difference version of model (3.10). The model includes interactions of the tax office dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.8: Heterogeneity in Response by Tax Office Type



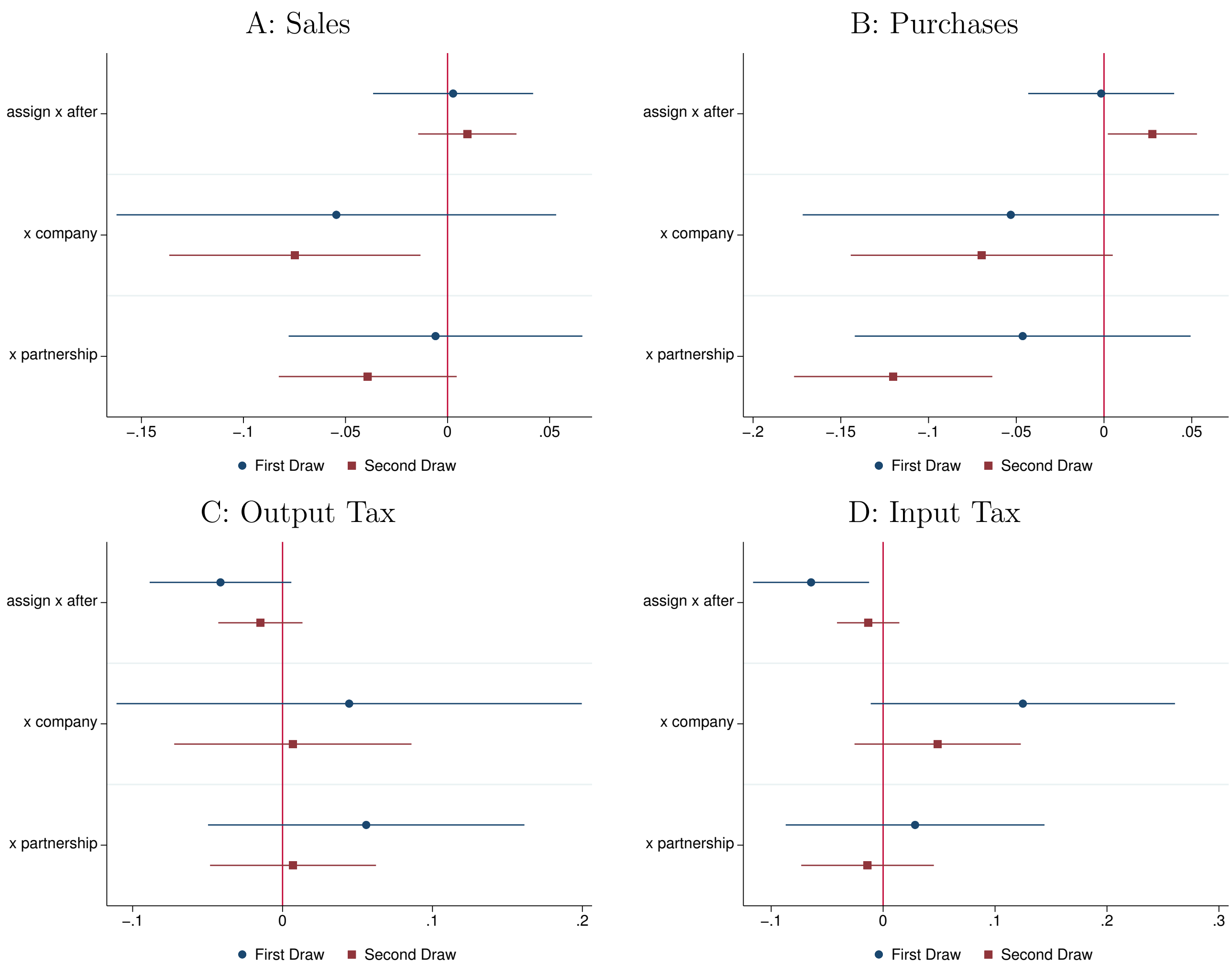
Notes: The figure explores heterogeneity in the audit effect. We divide firms into three groups based on the type of tax office they are subject to. Firms in four Large Taxpayer Units of the country are included in the first group (LTU), firms in the two Corporate Regional Tax Offices are included in the second group, and the rest of the firms are included in the baseline category. These firms are subject to a normal Regional Tax Office. We then estimate a triple-difference version of model (3.10). The model includes interactions of the tax office type dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.9: Heterogeneity in Response by Production Stage



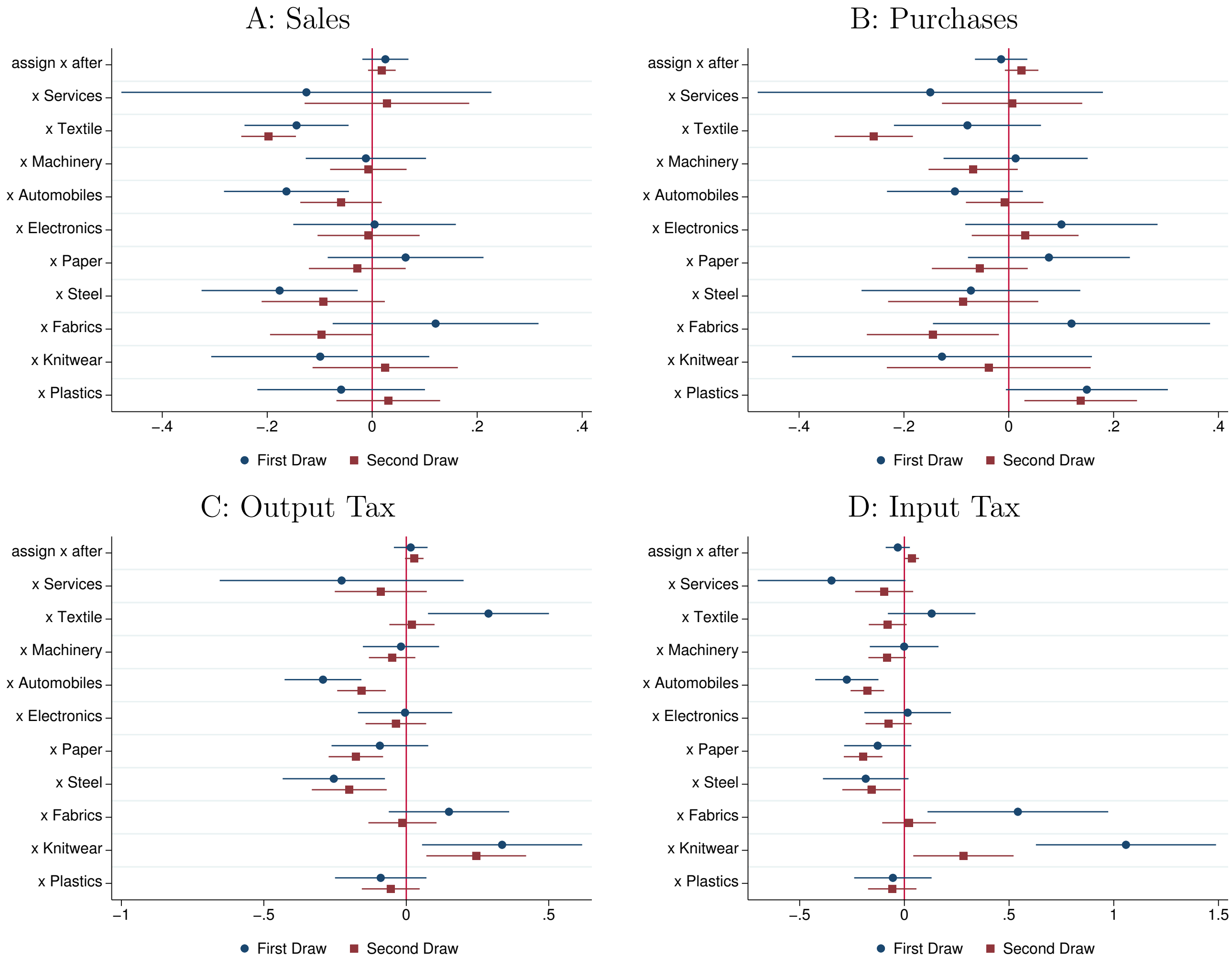
Notes: The figure explores heterogeneity in the audit effect. We divide firms into seven groups based on their principle business activity. The baseline category are retailers. These activities roughly capture the position of the firm in the supply chain. We then estimate a triple-difference version of model (3.10). The model includes interactions of the production stage dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.10: Heterogeneity in Response by Business Organization



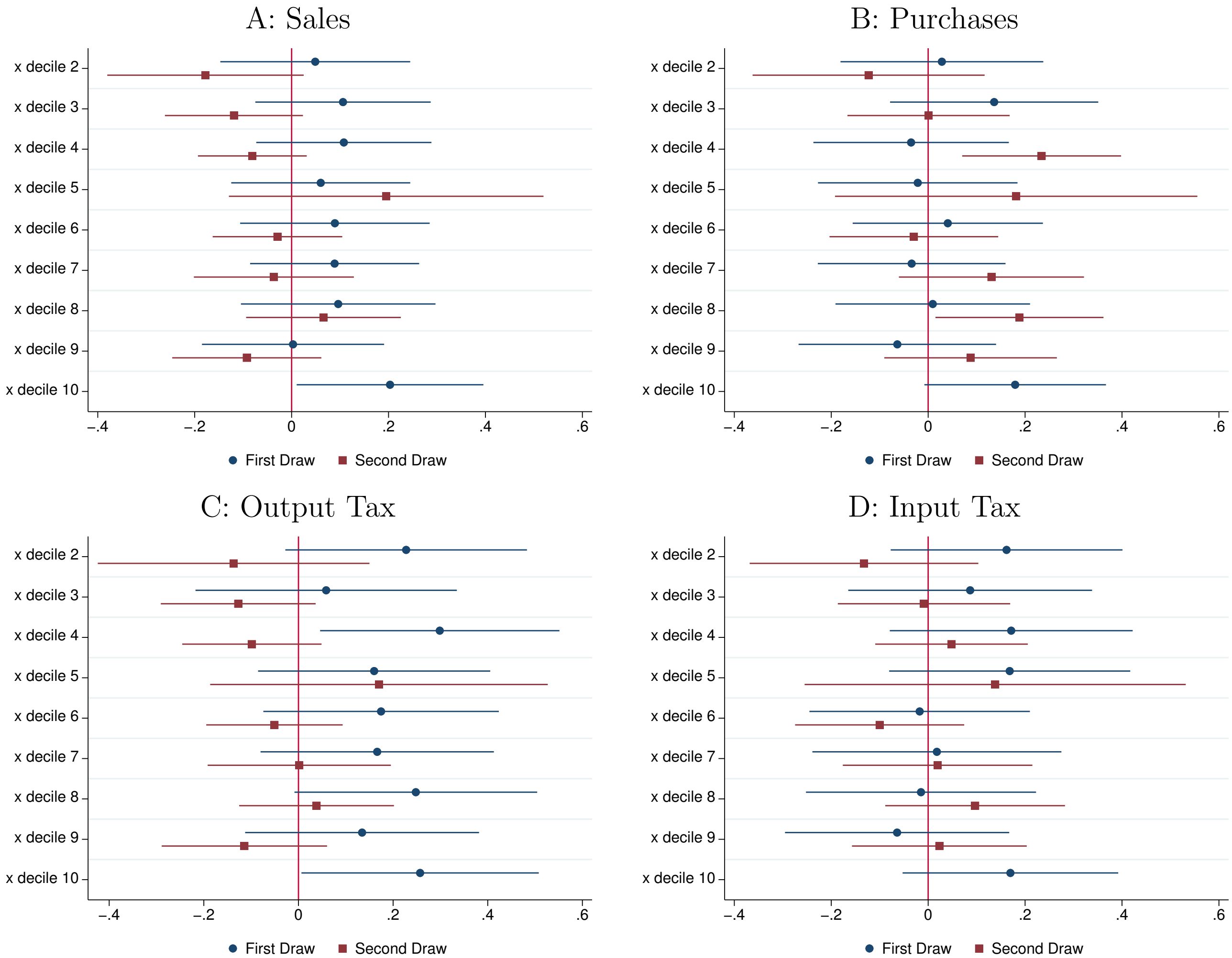
Notes: The figure explores heterogeneity in the audit effect. We divide firms into three groups based on their business organization. The baseline category are sole proprietors. We then estimate a triple-difference version of model (3.10). The model includes interactions of the business organization dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.11: Heterogeneity in Response by Industry



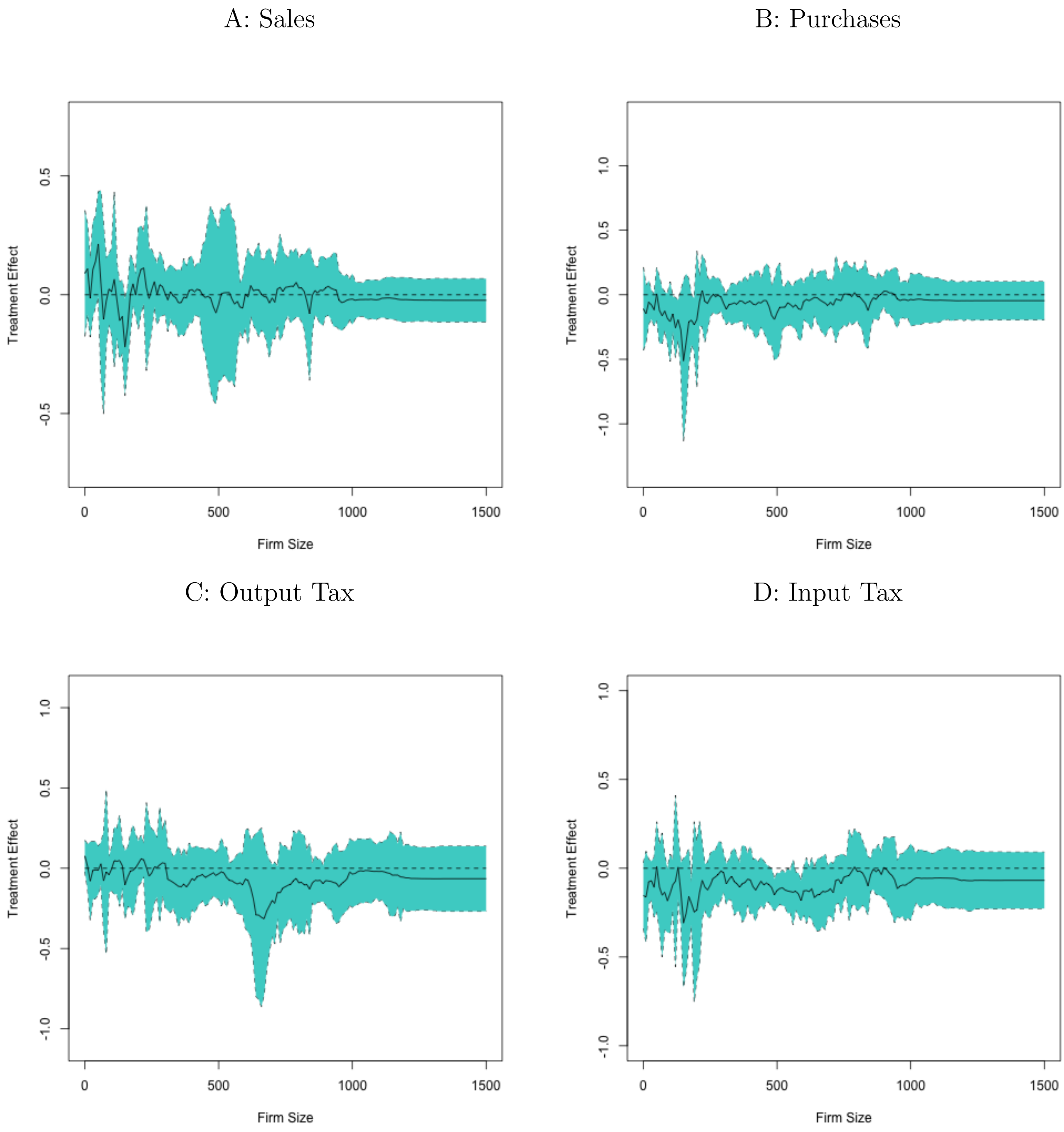
Notes: The figure explores heterogeneity in the audit effect. We divide firms into 12 groups based on the industry they operate in. We separate firms in 11 major industries of the country and club the rest into the baseline category. We then estimate a triple-difference version of model (3.10). The model includes interactions of the industry dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

Figure C.2.12: Heterogeneity in Response by Timing of Audit



Notes: The figure explores heterogeneity in the audit effect. We divide firms into ten deciles based on the time lag between the assignment and initiation of audit in days. We then estimate a triple-difference version of model (3.10). The model includes interactions of the firm decile dummies with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. We drop the triple-interaction term involving the first decile. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. The first wave results are in blue and the second wave results are in red. Standard errors are clustered at the firm level.

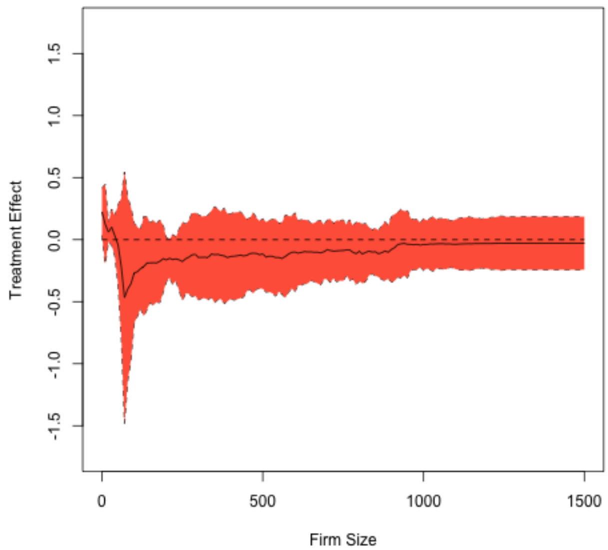
Figure C.2.13: Heterogeneity in Response by Firm Size (First Wave)



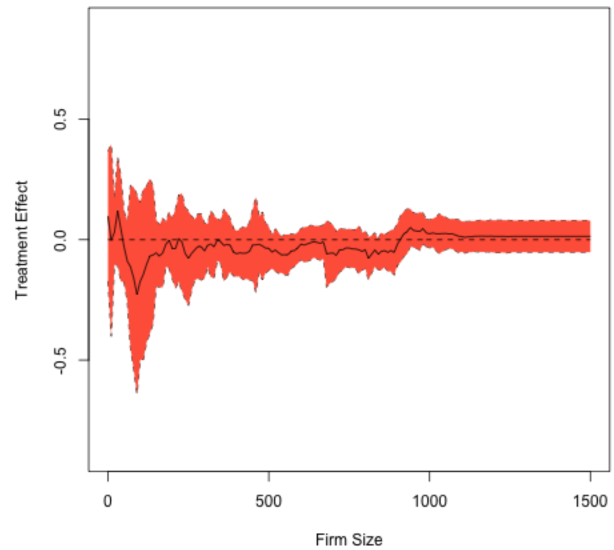
Notes: The figure explores heterogeneity in the audit effect. We use firm-size as a continuous variable. We then use a generalized random forest model to estimate the treatment effects of the audit for all values within the feasible range based on the available data. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The estimated treatment effects and 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for each outcome variable.

Figure C.2.14: Heterogeneity in Response by Firm Size (Second Wave)

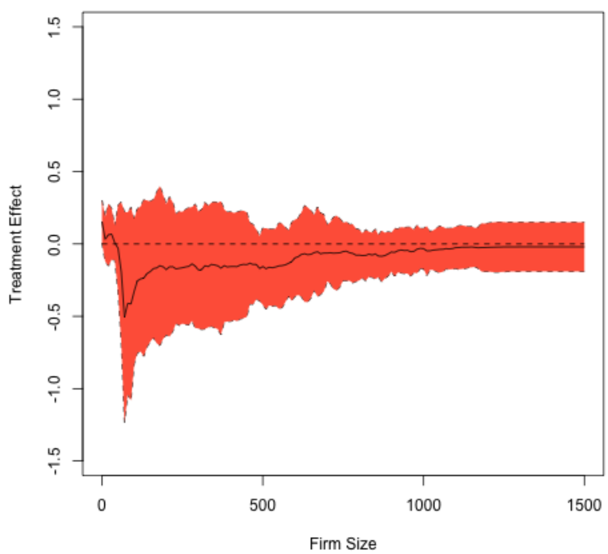
A: Sales



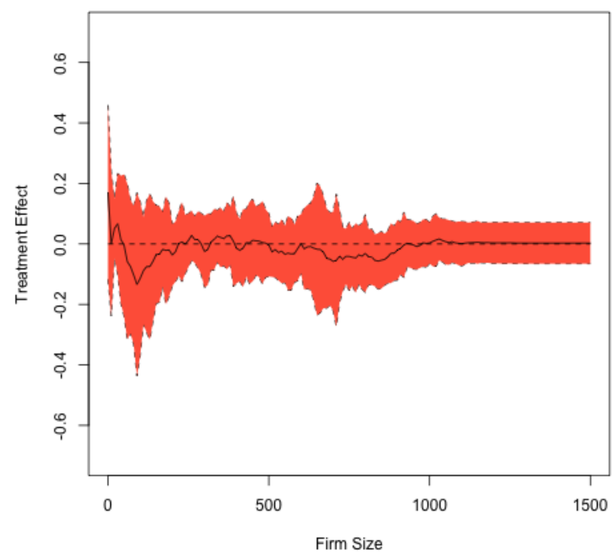
B: Purchases



C: Output Tax



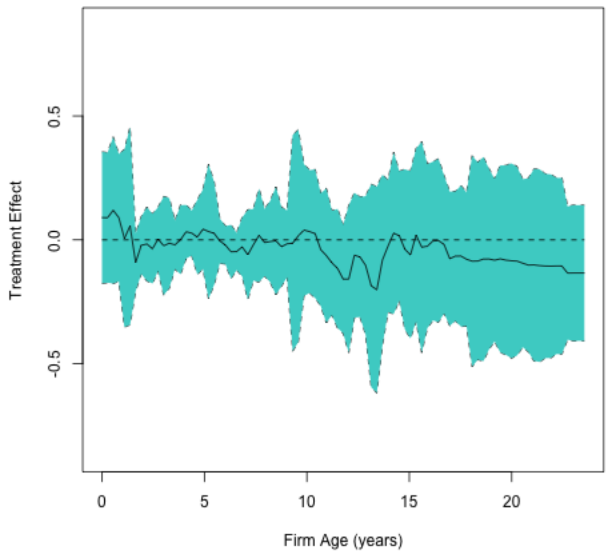
D: Input Tax



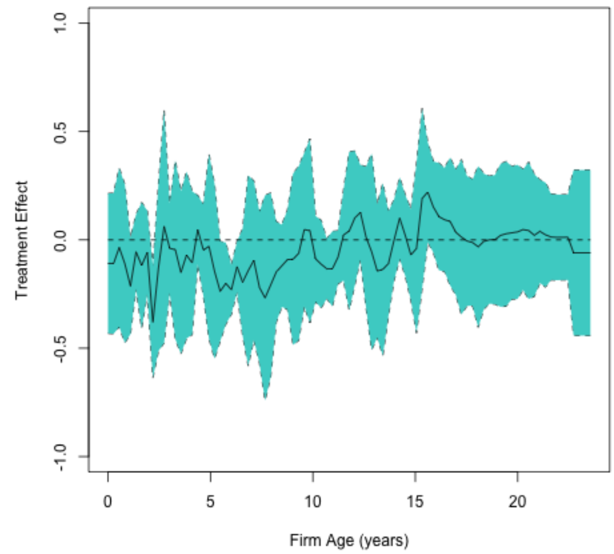
Notes: The figure explores heterogeneity in the audit effect. We use firm-size as a continuous variable. We then use a generalized random forest model to estimate the treatment effects of the audit for all values within the feasible range based on the available data. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The estimated treatment effects and 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for each outcome variable.

Figure C.2.15: Heterogeneity in Response by Firm Age (First Wave)

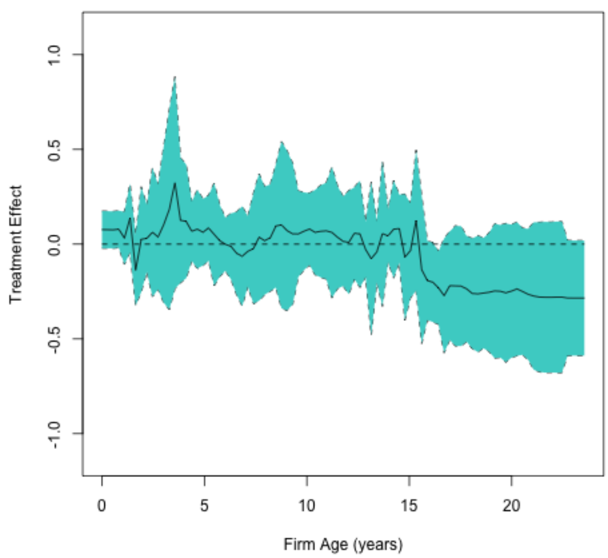
A: Sales



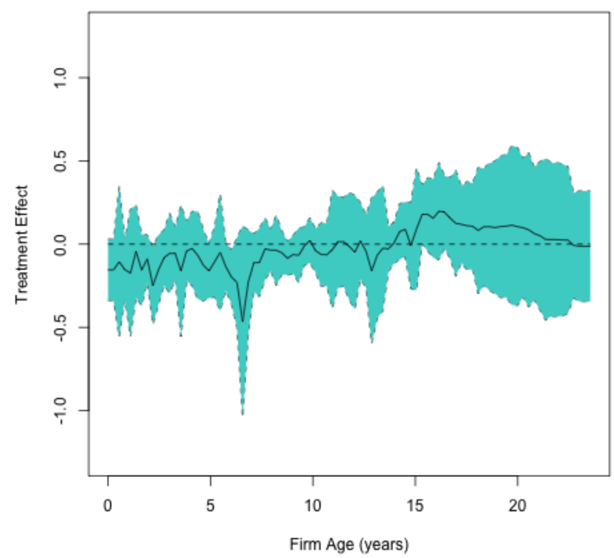
B: Purchases



C: Output Tax

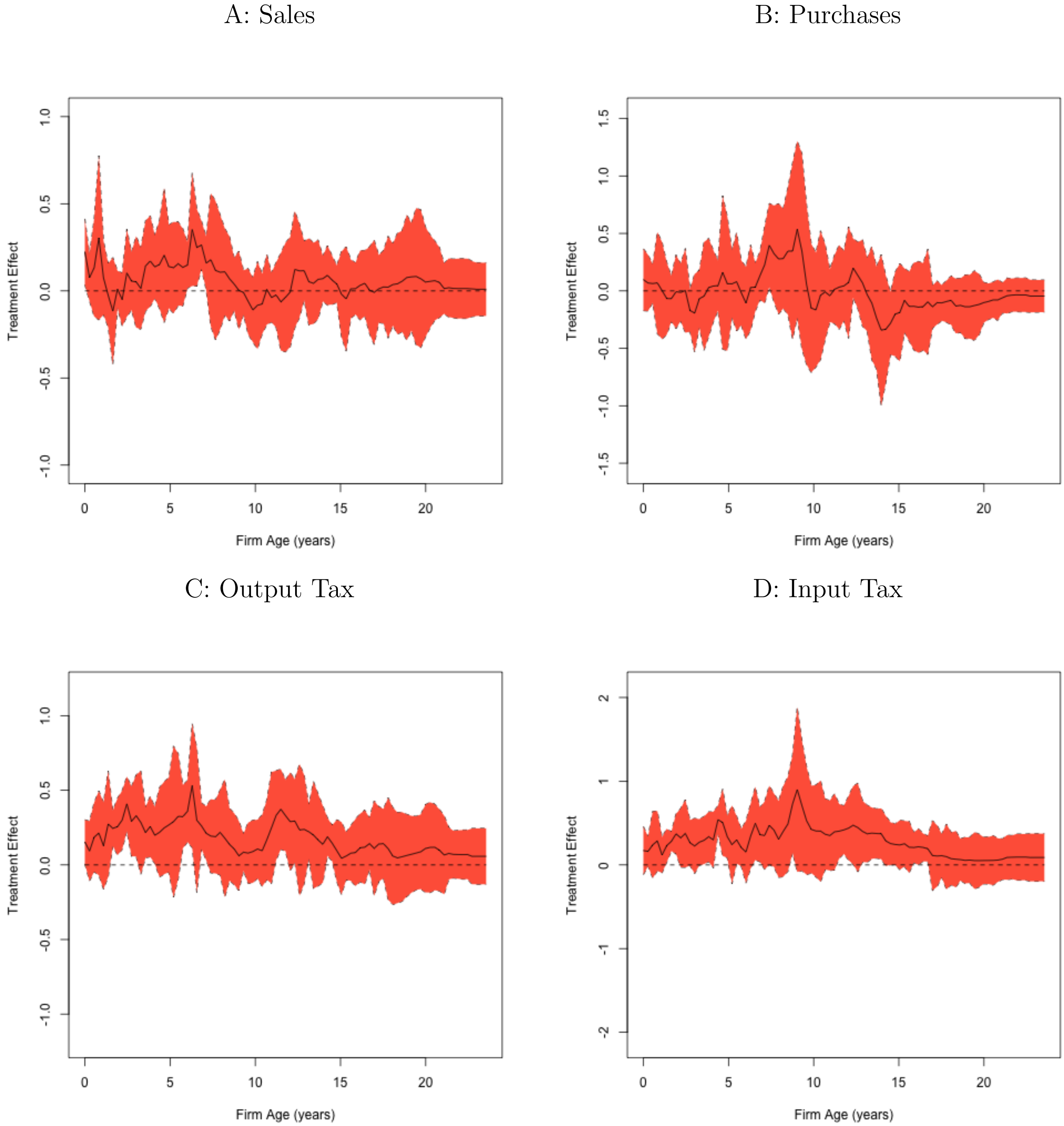


D: Input Tax



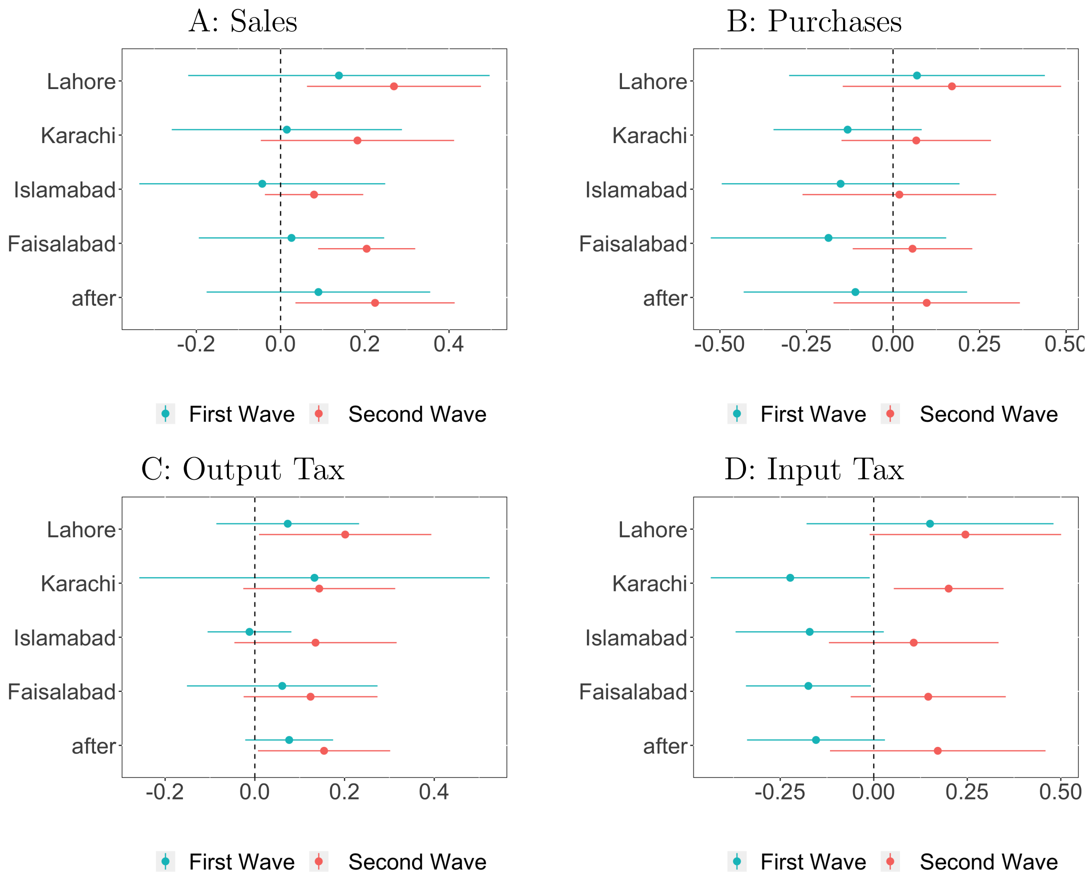
Notes: The figure explores heterogeneity in the audit effect. We use firm-age as a continuous variable. We then use a generalized random forest model to estimate the treatment effects of the audit for all values within the feasible range based on the available data. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The estimated treatment effects and 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for each outcome variable.

Figure C.2.16: Heterogeneity in Response by Firm Age (Second Wave)



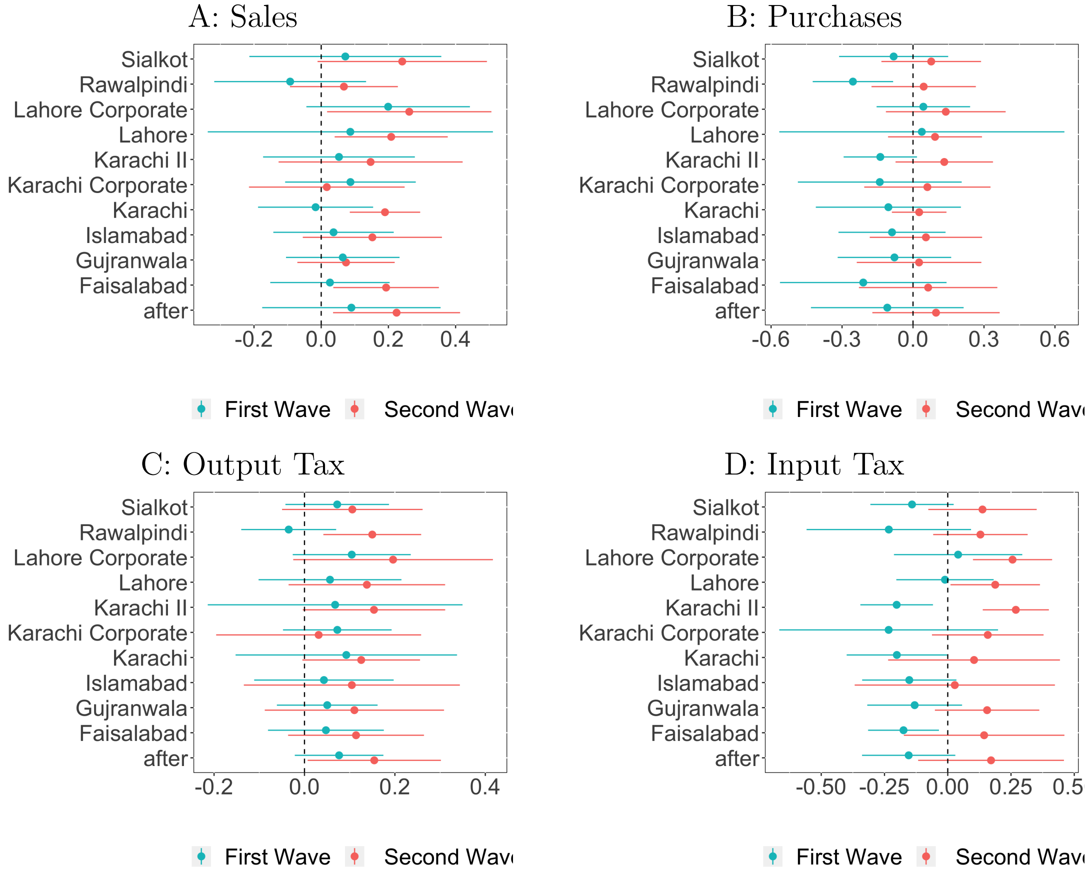
Notes: The figure explores heterogeneity in the audit effect. We use firm-age as a continuous variable. We then use a generalized random forest model to estimate the treatment effects of the audit for all values within the feasible range based on the available data. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The estimated treatment effects and 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for each outcome variable.

Figure C.2.17: Heterogeneity in Response by Location



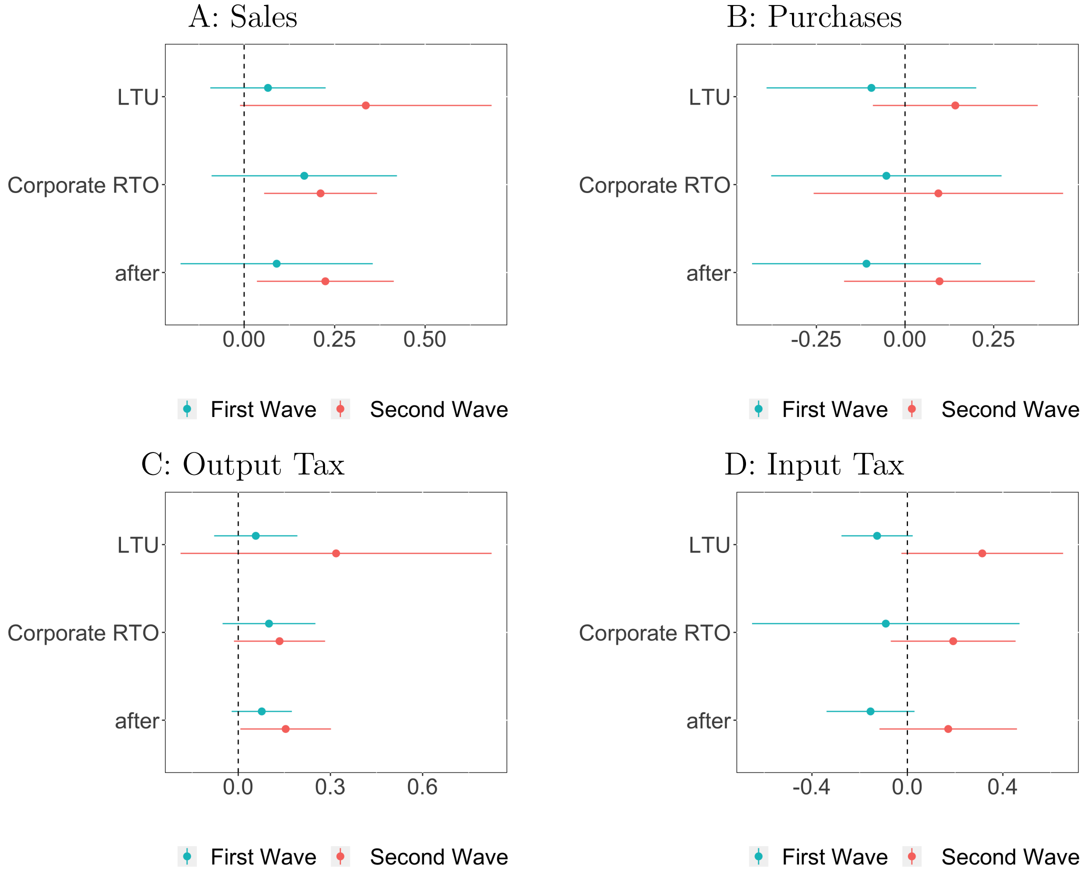
Notes: The figure explores heterogeneity in the audit effect. We divide firms into five groups depending upon the city their head office is located in. Firms not located in the four major cities of the country— Lahore, Karachi, Islamabad, and Faisalabad— are included in the baseline category. We then use a generalized random forest model to estimate the treatment effects of the audit. The model includes dummy variables for each group along with a dummy variable for "after" - indicating the time period after the date of the ballot. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The coefficients and the 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for the first and the second audit waves and for each outcome variable. The first wave results are in blue and the second wave results are in red.

Figure C.2.18: Heterogeneity in Response by Tax Office



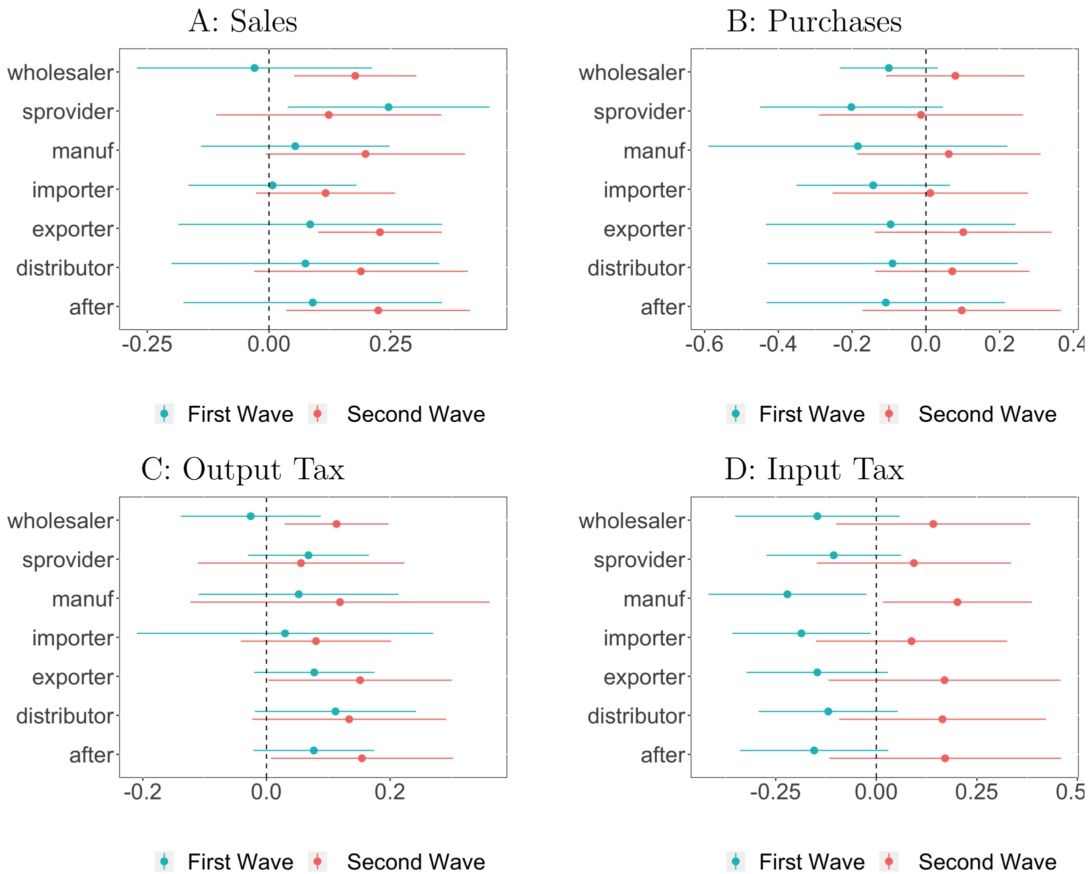
Notes: The figure explores heterogeneity in the audit effect. We divide firms into eleven groups based on the local tax office they are subject to. Firms not in the ten major tax offices are included in the baseline category. We then use a generalized random forest model to estimate the treatment effects of the audit. The model includes dummy variables for each group along with a dummy variable for "after" - indicating the time period after the date of the ballot. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The coefficients and the 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for the first and the second audit waves and for each outcome variable. The first wave results are in blue and the second wave results are in red.

Figure C.2.19: Heterogeneity in Response by Tax Office Type



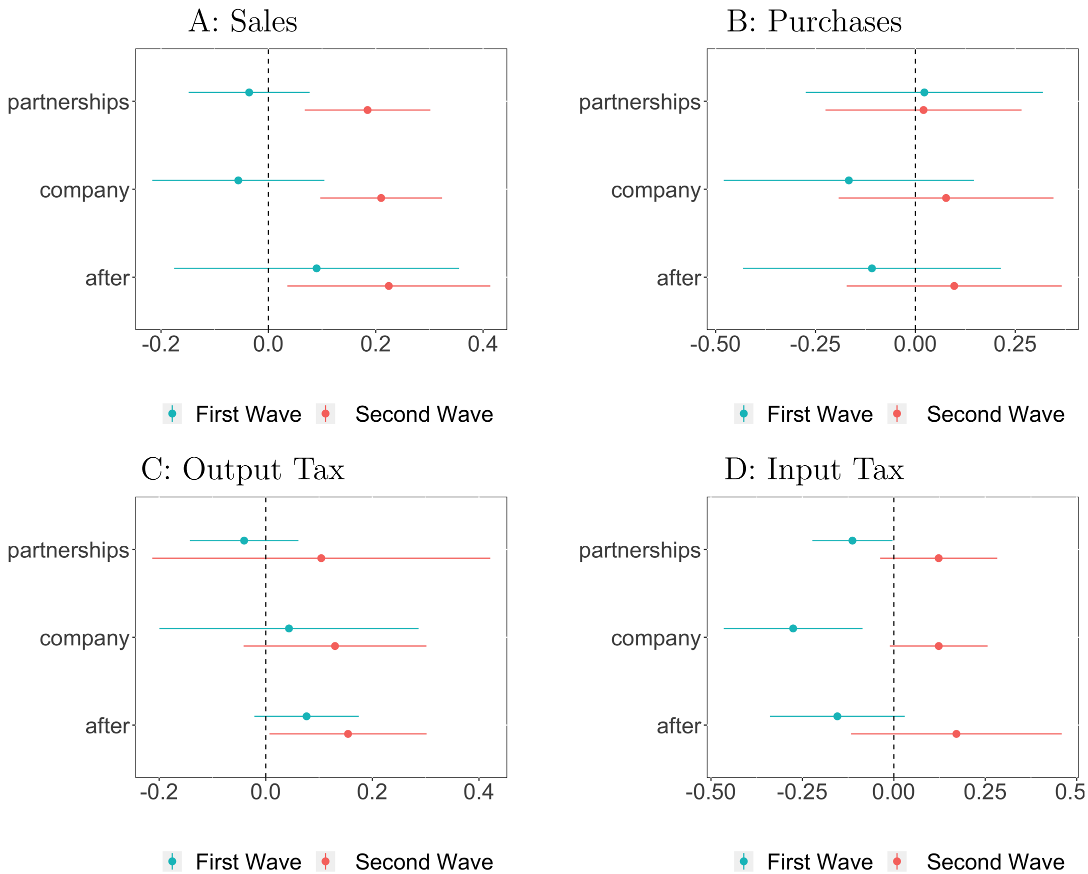
Notes: The figure explores heterogeneity in the audit effect. We divide firms into three groups based on the type of tax office they are subject to. Firms in four Large Taxpayer Units of the country are included in the first group (LTU), firms in the two Corporate Regional Tax Offices are included in the second group, and the rest of the firms are included in the baseline category. These firms are subject to a normal Regional Tax Office. We then use a generalized random forest model to estimate the treatment effects of the audit. The model includes dummy variables for each tax office type along with a dummy variable for "after" - indicating the time period after the date of the ballot. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The coefficients and the 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for the first and the second audit waves and for each outcome variable. The first wave results are in blue and the second wave results are in red.

Figure C.2.20: Heterogeneity in Response by Production Stage



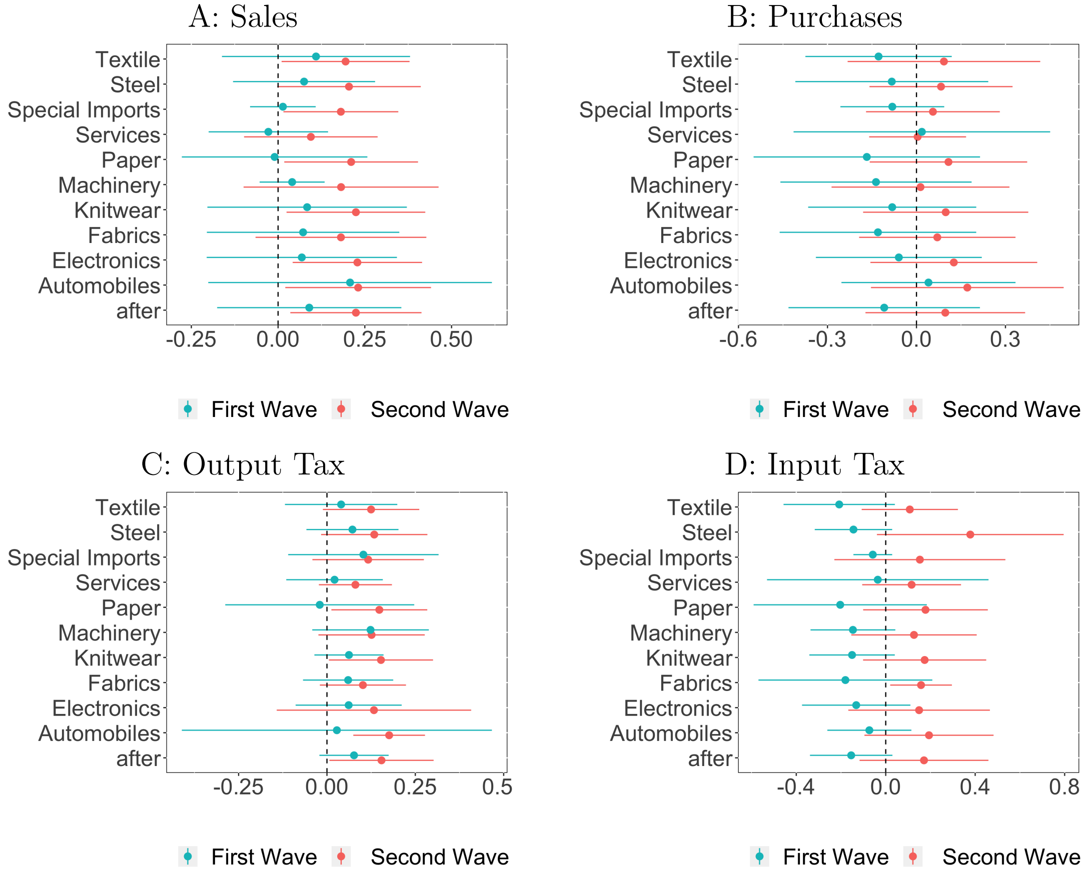
Notes: The figure explores heterogeneity in the audit effect. We divide firms into seven groups based on their principle business activity. The baseline category are retailers. These activities roughly capture the position of the firm in the supply chain. We then use a generalized random forest model to estimate the treatment effects of the audit. The model includes dummy variables for each group along with a dummy variable for "after" - indicating the time period after the date of the ballot. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The coefficients and the 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for the first and the second audit waves and for each outcome variable. The first wave results are in blue and the second wave results are in red.

Figure C.2.21: Heterogeneity in Response by Business Organization



Notes: The figure explores heterogeneity in the audit effect. We divide firms into three groups based on their business organization. The baseline category are sole proprietors We then use a generalized random forest model to estimate the treatment effects of the audit. The model includes dummy variables for each group along with a dummy variable for "after" - indicating the time period after the date of the ballot. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The coefficients and the 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for the first and the second audit waves and for each outcome variable. The first wave results are in blue and the second wave results are in red.

Figure C.2.22: Heterogeneity in Response by Industry



Notes: The figure explores heterogeneity in the audit effect. We divide firms into 11 groups based on the industry they operate in. We separate firms in 10 major industries of the country and club the rest into the baseline category. We then use a generalized random forest model to estimate the treatment effects of the audit. The model includes dummy variables for each group along with a dummy variable for "after" - indicating the time period after the date of the ballot. We consider a firm as treated (audited) if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The coefficients and the 95% confidence intervals on the estimated treatment effects are plotted. Models are estimated separately for the first and the second audit waves and for each outcome variable. The first wave results are in blue and the second wave results are in red.

Table C.2.1: Breakdown of the Detected Amount

	Amt. Detected		Amt. Recovered		Amt. Recoverable		Refund Curtailed	
	PKR	%	PKR	%	PKR	%	PKR	%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: First Audit Wave								
All Audited Firms	2.147	0.431	0.023	0.005	2.118	0.425	0.004	0.001
Amount Detected > 0	2.147	1.567	0.023	0.017	2.118	1.546	0.004	0.003
Size Quartile 1	0.062	684.756	0.001	11.221	0.061	673.534	0.000	0.000
Size Quartile 2	0.067	3.936	0.003	0.186	0.064	3.750	0.000	0.000
Size Quartile 3	0.215	1.746	0.008	0.067	0.203	1.648	0.003	0.021
Size Quartile 4	1.802	0.372	0.011	0.002	1.790	0.370	0.002	0.000
B: Second Audit Wave								
All Audited Firms	2.235	0.102	0.040	0.002	2.191	0.100	0.003	0.000
Amount Detected > 0	2.235	0.845	0.040	0.015	2.191	0.828	0.003	0.001
Size Quartile 1	0.045	10.205	0.002	0.473	0.042	9.649	0.000	0.000
Size Quartile 2	0.166	3.367	0.009	0.179	0.157	3.188	0.000	0.000
Size Quartile 3	0.217	0.889	0.009	0.036	0.205	0.840	0.003	0.012
Size Quartile 4	1.808	0.083	0.020	0.001	1.786	0.082	0.000	0.000

Notes*:* The table breaks down the total amount detected by audit (columns 1-2) into its three major components (columns 3-8). The odd-number columns report the amounts in PKR billions and the even-number columns the amount as a ratio of the aggregate annual turnover of the corresponding group of firm. Amount Recovered is the amount paid by the taxpayer as a result of audit. Amount Recoverable, on the other hand, is unpaid amount out of the total detected by audit. This amount is subject to quasi-judicial determination and appeal processes. Refund Curtailed indicates the amount by which the firm agreed to reduce its refund claim pending with the department.

Table C.2.2: Evasion Rate and Firm Characteristics

	Outcome: Tax Evasion Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Firm Size	-4.065***	-3.995***	-4.020***	-3.863***	-4.552***	-4.416***
	(0.481)	(0.450)	(0.454)	(0.457)	(0.484)	(0.516)
Share Manufacturers	0.091					0.004
	(0.282)					(0.281)
Share Retailers		0.142				0.182
		(0.176)				(0.175)
Share Major City			-0.116			-0.694
			(0.669)			(0.730)
Share Young Firms				0.783*		0.950*
				(0.447)		(0.489)
Share Textile					0.502***	0.503***
					(0.169)	(0.170)
Constant	100.526***	99.989***	102.595***	87.577***	105.866***	99.519***
	(7.466)	(7.483)	(12.735)	(10.578)	(7.575)	(13.414)
Observations	818	818	818	818	818	818

Notes*:* The table breaks down the total amount detected by audit (columns 1-2) into its three major components (columns 3-8). The odd-number columns report the amounts in PKR billions and the even-number columns the amount as a ratio of the aggregate annual turnover of the corresponding group of firm. Amount Recovered is the amount paid by the taxpayer as a result of audit. Amount Recoverable, on the other hand, is unpaid amount out of the total detected by audit. This amount is subject to quasi-judicial determination and appeal processes. Refund Curtailed indicates the amount by which the firm agreed to reduce its refund claim pending with the department.

Table C.2.3: Selection in Sequencing of Audits

	Outcome: Days between assignment and initiation			
	(1)	(2)	(3)	(4)
Sales	-1.785	-4.301	2.542	2.679
	(7.492)	(7.489)	(2.749)	(2.657)
Purchases	-0.727	-3.636	-2.583	0.588
	(8.569)	(8.568)	(5.626)	(5.433)
Output Tax	9.936	8.030	-2.929	0.718
	(30.624)	(30.012)	(12.057)	(11.651)
Input Tax	-4.050	1.030	3.229	-2.713
	(14.118)	(13.982)	(11.034)	(10.648)
Tax Paid	-6.673	-3.513	-1.108	-2.919
	(23.011)	(22.538)	(4.718)	(4.554)
Exports	-0.550	-0.126	1.836	2.399
	(1.560)	(1.540)	(1.002)	(0.974)
Imports	-0.201	-0.223	-0.370	-0.264
	(1.884)	(1.916)	(0.643)	(0.624)
Refund	1.382	1.662	-1.847	-2.325
	(1.395)	(1.377)	(0.866)	(0.840)
Carry Forward	1.734	1.132	-0.143	-0.300
	(3.374)	(3.355)	(0.569)	(0.549)
Manufacturer	-13.271	-11.003	-1.860	-1.986
	(5.331)	(5.298)	(1.615)	(1.581)
Importer	-0.785	-0.614	-3.302	0.310
	(6.230)	(6.190)	(1.833)	(1.791)
Exporter	1.834	6.001	-1.649	-1.134
	(9.390)	(9.301)	(2.282)	(2.295)
Distributor	7.098	9.746	-0.251	-1.645
	(9.143)	(8.977)	(2.469)	(2.395)
Wholesaler	-5.548	-2.847	-1.848	0.958
	(5.391)	(5.315)	(1.669)	(1.624)
Service Provider	-7.959	-4.111	0.109	1.141
	(5.332)	(5.247)	(1.661)	(1.606)
Constant	46.995	44.436	18.961	17.830
	(4.843)	(4.768)	(1.490)	(1.443)
Observations	3,482	3,481	3,612	3,611
Corporation FEs	Yes	Yes	Yes	Yes
Tax Office FEs	No	Yes	No	Yes

Notes*:* The table explores selection in audit. We regress the time lag measured in number of days between the assignment and initiation of audit on baseline firm characteristics. We standardize the first nine variables in this table by subtracting the mean and dividing by the standard deviation of the variable. Since audits were taken up by local tax offices, we include the tax office fixed effects in even-numbered columns. The first two columns report results for the first audit wave and the last two for the second audit wave. Standard errors are in parenthesis.

Table C.2.4: Preexisting Trends

	First Wave					Second Wave				
	Sales	Purchases	Output	Input	Tax Payable	Sales	Purchases	Output	Input	Tax Payable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
assign $\times$ year $\in [s-1, s]$	-0.018	-0.005	-0.039	-0.004	-0.033	-0.016	-0.018	-0.027	-0.030	0.002
	(0.015)	(0.017)	(0.020)	(0.021)	(0.025)	(0.010)	(0.011)	(0.013)	(0.014)	(0.017)
assign $\times$ year $\in [s-3, s]$	0.001	0.021	-0.031	0.021	-0.006	-0.006	-0.014	-0.005	-0.020	0.012
	(0.014)	(0.016)	(0.018)	(0.020)	(0.022)	(0.010)	(0.012)	(0.014)	(0.014)	(0.017)
assign $\times$ year $\in [s-5, s]$	-0.004	0.042	-0.019	0.040	0.051	0.028	0.007	0.020	-0.002	0.056
	(0.021)	(0.022)	(0.022)	(0.024)	(0.033)	(0.011)	(0.012)	(0.013)	(0.014)	(0.017)
Observations	2,324,186	2,025,380	1,672,095	1,681,583	1,154,574	2,628,878	2,290,848	1,934,273	1,945,733	1,312,928
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table explores if the preexisting trends for the five outcomes indicated in the heading of each column were parallel between firms who were picked for audit in a random ballot and other firms in the eligible sample. We estimate a model similar to (3.10) replacing the $assign \times after_{it}$ dummy with three dummies shown in the top three rows. The dummy variable $assign_i$ denotes that firm i's was picked for audit in the random ballot indicated in the heading of the column. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The sample for these regressions include the baseline periods only, from July 2008 to August 2013 for the first wave and from July 2008 to August 2014 for the second. The dummy variable $year \in [s-1, s]$ indicates that the period is one of the last twelve months included in the regression and so on. Standard errors are in parenthesis, which have been clustered at the firm level.

Table C.2.5: Impacts of Random Audits Assigned in the First Wave

	Impacts After One Year					Impacts After Three Years				
	Sales	Purchases	Output Tax	Input Tax	Tax Payable	Sales	Purchases	Output Tax	Input Tax	Tax Payable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A: ITT Estimates										
assign × after	-0.009	-0.009	-0.016	-0.017	-0.037	-0.007	-0.021	-0.025	-0.036	-0.015
	(0.016)	(0.019)	(0.021)	(0.026)	(0.027)	(0.014)	(0.019)	(0.019)	(0.023)	(0.030)
Observations	2,802,387	2,456,864	2,061,472	2,089,489	1,393,541	3,809,614	3,315,994	2,857,885	2,895,330	1,890,220
B: LATE Estimates										
treat × after	-0.013	-0.014	-0.022	-0.024	-0.051	-0.010	-0.030	-0.035	-0.051	-0.021
	(0.022)	(0.027)	(0.029)	(0.037)	(0.036)	(0.019)	(0.027)	(0.026)	(0.031)	(0.041)
Observations	2,802,387	2,456,864	2,061,472	2,089,489	1,393,541	3,809,614	3,315,994	2,857,885	2,895,330	1,890,220
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table estimates the impact of audit on firms' future behavior. In the top panel, the coefficient assign × after shows $\hat{\gamma}$ from model (3.10), where the dummy variable assign$_i$ denotes that firm i's audit was assigned through the first random ballot held on September 13, 2013. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable after$_t$ indicates that month t falls after the date of the ballot. The sample includes periods up to October 2014 for the first five columns and periods up to October 2016 for the rest. Panel B shows the corresponding results from 2sls regressions, where the endogenous variable audit$_i$ is instrumented by the initial random assignment. Standard errors are in parenthesis, which have been clustered at the tax office level.

Table C.2.6: Audit Impacts – First Stage

Outcome:	$audit \times after_{it}$					
Random Draw Held On:	September 13, 2013		September 25, 2014		September 14, 2015	
Post Sample:	One Year	Three Years	One Year	Three Years	One Year	Three Years
	(1)	(2)	(3)	(4)	(5)	(6)
assign × after	0.704	0.703	0.294	0.296	0.133	0.134
	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	6,893,186	9,681,146	7,894,004	10721371	8,241,185	10829729
F Statistic	10,353	10,071	4,751	4,658	1,120	1,102

Notes: The table reports the first stage of our 2sls models. We estimate model (3.10) using the dummy $treat \times after_{it}$ as the outcome variable, where $treat_i$ takes the value 1 if firm i was audited in the corresponding audit wave indicated in the heading of each column. The coefficient assign × after shows $\hat{\gamma}$ from these regressions. The dummy variable assign_i denotes that firm i's audit was assigned through the random ballot indicated in the heading of each column. The sample includes the population of VAT filers excluding government departments and firms already under audit. The dummy variable after_t indicates that month t falls after the date of the ballot. We report results for two Post Samples: One Year specifications include twelve $after_t$ periods and Three Years specifications include 36 $after_t$ periods. In each case, the samples includes all months from July 2008 to the last $after_t$ period. Standard errors are in parenthesis, which have been clustered at the firm level.

Table C.2.7: Impacts of Random Audits Assigned in the Third Wave

	Impacts After One Year					Impacts After Three Years				
	Sales	Purchases	Output Tax	Input Tax	Tax Payable	Sales	Purchases	Output Tax	Input Tax	Tax Payable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
assign × after	-0.034	-0.024	-0.039	-0.009	0.004	-0.050	-0.040	-0.071	-0.076	-0.093
	(0.011)	(0.013)	(0.014)	(0.014)	(0.014)	(0.011)	(0.014)	(0.015)	(0.015)	(0.015)
Observations	3,007,568	2,590,734	2,256,294	2,265,080	2,758,303	3,910,133	3,341,025	2,879,242	2,930,477	3,577,794
B: LATE Estimates										
treat × after	-0.261	-0.185	-0.296	-0.063	0.033	-0.376	-0.297	-0.487	-0.527	-0.652
	(0.083)	(0.102)	(0.106)	(0.105)	(0.108)	(0.087)	(0.106)	(0.110)	(0.108)	(0.112)
Observations	3,007,568	2,590,734	2,256,294	2,265,080	2,758,303	3,910,133	3,341,025	2,879,242	2,930,477	3,577,794
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table estimates the impact of audit on firms' future behavior. In the top panel, the coefficient assign × after shows $\hat{\gamma}$ from model (3.10), where the dummy variable assign_i denotes that firm i's audit was assigned through the first random ballot held on September 14, 2015. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable after_t indicates that month t falls after the date of the ballot. The sample includes periods up to October 2016 for the first five columns and periods up to October 2018 for the rest. Panel B shows the corresponding results from 2sls regressions, where the endogenous variable audit_i is instrumented by the initial random assignment. Standard errors are in parenthesis, which have been clustered at the firm level.

Table C.2.8: Preexisting Trends – Audited vs. Not Audited

	First Wave					Second Wave				
	Sales	Purchases	Output	Input	Tax Payable	Sales	Purchases	Output	Input	Tax Payable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
treat $\times$ year $\in [s-1, s]$	0.019	0.038	-0.016	0.022	-0.046	0.001	0.020	0.022	-0.024	-0.007
	(0.016)	(0.018)	(0.021)	(0.020)	(0.026)	(0.015)	(0.019)	(0.023)	(0.022)	(0.028)
treat $\times$ year $\in [s-3, s]$	0.070	0.074	0.006	0.071	0.029	0.003	0.011	0.037	0.029	-0.006
	(0.016)	(0.017)	(0.019)	(0.019)	(0.024)	(0.015)	(0.019)	(0.023)	(0.022)	(0.027)
treat $\times$ year $\in [s-5, s]$	0.089	0.066	0.011	0.066	0.098	0.034	0.028	0.054	0.064	0.025
	(0.024)	(0.022)	(0.022)	(0.024)	(0.033)	(0.018)	(0.019)	(0.022)	(0.022)	(0.028)
Observations	2,324,186	2,025,380	1,672,095	1,681,583	1,154,574	2,628,878	2,290,848	1,934,273	1,945,733	1,312,928
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table explores if the preexisting trends for the five outcomes indicated in the heading of each column were parallel between audited and unaudited firms. We estimate a model similar to (3.10) replacing the $assign \times after_{it}$ dummy with three dummies shown in the top three rows. The dummy variable $treat_i$ denotes that firm i was audited in the wave indicated in the heading of the column. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The sample for these regressions include the baseline periods only, from July 2008 to August 2013 for the first wave and from July 2008 to August 2014 for the second. The dummy variable $year \in [s-1, s]$ indicates that the period is one of the last twelve months included in the regression and so on. Standard errors are in parenthesis, which have been clustered at the firm level.

Table C.2.9: Heterogeneity in Response With Respect To Amount Detected

	Sales	Purchases	Output Tax	Input Tax	Tax Payable
	(1)	(2)	(3)	(4)	(5)
A: First Wave					
assign × after	-0.009	-0.016	-0.020	-0.029	0.004
	(0.019)	(0.021)	(0.025)	(0.026)	(0.031)
assign × after × trait	0.009	-0.023	-0.022	-0.031	-0.089
	(0.040)	(0.048)	(0.052)	(0.054)	(0.070)
Observations	3,839,502	3,328,628	2,884,225	2,906,045	1,913,096
B: Second Wave					
assign × after	-0.014	-0.019	-0.016	-0.009	0.005
	(0.011)	(0.013)	(0.013)	(0.013)	(0.017)
assign × after × trait	0.040	0.119	0.053	0.038	0.010
	(0.031)	(0.041)	(0.042)	(0.039)	(0.048)
Observations	4,390,478	3,791,277	3,262,221	3,313,664	2,151,912
Firm FEs	Yes	Yes	Yes	Yes	Yes
Period FEs	Yes	Yes	Yes	Yes	Yes

Notes: The table explores heterogeneity in the audit effect. We divide firms into two groups. Firms against whom a positive amount was detected by audit are included in one group (indicated by the dummy variable $trait_i$); the rest of the firms are included in the baseline category. We then estimate a triple-difference version of model (3.10). The model includes interactions of the $trait_i$ dummy with the $assign \times after_{it}$ dummy. The $assign_i$ dummy takes the value 1 if the firm's audit was assigned in the corresponding random computer ballot. The control group comprises the rest of the firms in the eligible sample. The eligible sample consists of the population of VAT filers excluding government departments and firms already under audit. The dummy variable $after_t$ indicates that month t falls after the date of the ballot. The coefficients and the 95% confidence intervals on the double and triple-interaction terms from these regressions are plotted. Regressions are run separately for the first and the second audit waves. Standard errors are clustered at the firm level.

Table C.2.10: Heterogeneity in Amount Detected by Share Final Sales

	Outcome: Amount Detected (Std. Deviations)					
	(1)	(2)	(3)	(4)	(5)	(6)
A: Share Final Sales						
2nd Quartile	-0.100*	-0.099**	-0.097*	-0.098*	-0.105**	-0.096**
	(0.051)	(0.051)	(0.050)	(0.050)	(0.052)	(0.048)
3rd Quartile	-0.094*	-0.091*	-0.085*	-0.090*	-0.098*	-0.086*
	(0.051)	(0.050)	(0.047)	(0.050)	(0.052)	(0.046)
4th Quartile	-0.101**	-0.097*	-0.090**	-0.085*	-0.108*	-0.085*
	(0.051)	(0.049)	(0.046)	(0.044)	(0.056)	(0.045)
Observations	6,561	6,561	6,561	6,560	6,548	6,547
B: Share (Final Sales + Purchases from Unregistered Sector)						
2nd Quartile	-0.085	-0.082	-0.076	-0.081	-0.088	-0.074
	(0.052)	(0.051)	(0.048)	(0.051)	(0.053)	(0.046)
3rd Quartile	-0.108**	-0.087*	-0.094**	-0.083*	-0.113**	-0.074*
	(0.052)	(0.045)	(0.045)	(0.043)	(0.057)	(0.042)
4th Quartile	-0.113**	-0.086**	-0.095**	-0.086**	-0.118**	-0.076*
	(0.052)	(0.044)	(0.043)	(0.043)	(0.059)	(0.043)
Observations	6,561	6,561	6,561	6,560	6,548	6,547
Size FEs	No	Yes	No	No	No	Yes
Production Stage FEs	No	No	Yes	No	No	Yes
Tax Office FEs	No	No	No	Yes	No	Yes
Industry FEs	No	No	No	No	Yes	Yes

Notes*:* The table examines if the amount detected by audit changes with the share of final sales reported by a firm at the baseline. The outcome variable here is the amount detected by audit, normalized by its standard deviation. To maximize power, we pool together the audits conducted in the first two waves. Final sales are defined as sales where the other party to the transaction does not possess a national tax number: they are either consumers or informal firms. We divide firms into four quartiles based on the share of final sales in their turnover at the baseline. We regress the outcome variable on the three quartile dummies, omitting the first quartile as the reference group. We successively introduce the controls indicated in the last four lines. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Bibliography

Abadie, Alberto (Apr. 2003). "Semiparametric instrumental variable estimation of treatment response models". In: *Journal of Econometrics* 113.2, pp. 231–263. DOI: 10.1016/s0304-4076(02)00201-4.

Advani, Arun, William Elming, and Jonathan Shaw (2019). *The dynamic effects of tax audits.*

Agrawal, David and Mohammed Mardan (2019). "Will Destination-based Taxes Be Fully Exploited When Available? An Application to the U.S. Commodity Tax System". In: *Journal of Public Economics* 169, pp. 128–143.

Agrawal, David R and Laura Zimmerman (2019). *Production and evasion responses with limited state capacity: Evidence from major tax reforms in India*. Tech. rep. IGC Working Paper S-89411-INC-1.

Allingham, Michael G. and Agnar Sandmo (1972). "Income Tax Evasion: A Theoretical Analysis". In: *Journal of Public Economics*, pp. 323–338.

Almunia, Miguel and David Lopez-Rodriguez (Feb. 2018). "Under the Radar: The Effects of Monitoring Firms on Tax Compliance". In: *American Economic Journal: Economic Policy* 10.1, pp. 1–38. DOI: 10.1257/pol.20160229.

Antinyan, Armenak and Zareh Asatryan (2019). "Nudging for Tax Compliance: A Meta-Analysis". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3500744.

Athey, Susan, Julie Tibshirani, and Stefan Wager (Apr. 2019). "Generalized random forests". In: *The Annals of Statistics* 47.2. DOI: 10.1214/18-aos1709.

Bank, World (2019). *Doing Business 2020*. Washington, D.C: World Bank Publications. ISBN: 9781464814419.

Basri, M. Chatib et al. (Aug. 2019). *Tax Administration vs. Tax Rates: Evidence from Corporate Taxation in Indonesia*. Tech. rep. DOI: 10.3386/w26150.

Bénabou, Roland and Jean Tirole (Aug. 2016). "Mindful Economics: The Production, Consumption, and Value of Beliefs". In: *Journal of Economic Perspectives* 30.3, pp. 141–164. DOI: 10.1257/jep.30.3.141.

Bérgolo, Marcelo et al. (July 2017). *Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment.* Tech. rep. DOI: 10.3386/w23631.

Besley, Timothy and Torsten Persson (2013). "Taxation and Development". In: *handbook of public economics*. Ed. by Alan J. Auerbach et al. Vol. 5.

— (Nov. 2014). "Why Do Developing Countries Tax So Little?" In: *Journal of Economic Perspectives* 28.4, pp. 99–120. DOI: 10.1257/jep.28.4.99.

Best, Michael et al. (2015). "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan". In: *Journal of Political Economy* 123.6, pp. 1311–1356.

Bickley, James M. (2003). *Value Added Tax: Concepts, Policy Issues, and OECD Experiences.* Ed. by Susan Boriotti and Donna Dennis. Hauppauge, New York, United States of America: Novinka Books.

Bird, Richard M. and Pierre-Pascal Gendron (Dec. 2000). "CVAT, VIVAT, and Dual VAT: Vertical "Sharing" and Interstate Trade". In: *International Tax and Public Finance* 7.6, pp. 753–761.

— (2007). *The VAT in Developing and Transitional Countries.* Cambridge: Cambridge University Press.

Blomquist, Soren et al. (Mar. 2021). "On Bunching and Identification of the Taxable Income Elasticity". In: *Journal of Political Economy.* DOI: 10.1086/714446.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall (Aug. 2017). "Beyond LATE with a Discrete Instrument". In: *Journal of Political Economy* 125.4, pp. 985–1039. DOI: 10.1086/692712.

Card, David et al. (2015). "Inference on Causal Effects in a Generalized Regression Kink Design". In: *Econometrica* 83.6, pp. 2453–2483. DOI: 10.3982/ecta11224.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal (2017). "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement". In: *American Economic Journal: Applied Economics*, pp. 144–164.

Chetty, Raj et al. (May 2011). "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records". In: *The quarterly journal of economics* 126, pp. 749–804. DOI: 10.1093/qje/qjr013.

Crawford, Ian, Michael Keen, and Stephen Smith (2010). "Value Added Tax and Excises". In: *Dimensions of Tax Design: the Mirrlees Review.* Ed. by Stuart Adam et al. Oxford: Oxford University Press, pp. 275–422.

Davoodi, Hamid R. and Vito Tanzi (2000). *Corruption, Growth and Public Finances.* Research rep. wp/00/182. Washington D. C.

Deaton, Angus and Nancy Cartwright (Aug. 2018). "Understanding and misunderstanding randomized controlled trials". In: *Social Science & Medicine* 210, pp. 2–21. DOI: 10.1016/j.socscimed.2017.12.005.

DeBacker, Jason et al. (May 2015). "Legal Enforcement and Corporate Behavior: An Analysis of Tax Aggressiveness after an Audit". In: *The Journal of Law and Economics* 58.2, pp. 291–324. DOI: 10.1086/684037.

— (Feb. 2018). "Once Bitten, Twice Shy? The Lasting Impact of Enforcement on Tax Compliance". In: *The Journal of Law and Economics* 61.1, pp. 1–35. DOI: 10.1086/697683.

Duflo, Esther (May 2017). "The Economist as Plumber". In: *American Economic Review* 107.5, pp. 1–26. DOI: 10.1257/aer.p20171153.

Emran, M. Shahe and Joseph E. Stiglitz (2005). "On selective indirect tax reform in developing countries". In: *Journal of Public Economics* 89, pp. 599–623.

Fan, Haichao et al. (Mar. 2018). "The Dynamic Effects Of Computerized VAT Invoices On Chinese Manufacturing Firms". Cambridge, MA, USA. URL: http://www.nber.org/papers/w24414.

Fath, Meredith Stevenson, Robert Goulder, and Risa Williams, eds. (2015). *(2011)*. Tax Analysts. URL: http://www.taxanalysts.com/www/freefiles.nsf/Files/VATReader.pdf/%5C$file/VATReader.pdf.

FBR (2015). *Federal Board of Revenue, Pakistan, Taxpayer's Audit Wing.*

Gadenne, Lucie, Tushar Nandi, and Roland Rathelot (2019). *Taxation and Supplier Networks: Evidence from India.* Mimeo.

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks (Jan. 2020). "Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test". In: *American Economic Journal: Applied Economics* 12.1, pp. 1–31. DOI: 10.1257/app.20170717.

Gemmell, Norman and Marisa Ratto (Mar. 2012). "BEHAVIORAL RESPONSES TO TAXPAYER AUDITS: EVIDENCE FROM RANDOM TAXPAYER INQUIRIES". In: *National Tax Journal* 65.1, pp. 33–57. DOI: 10.17310/ntj.2012.1.02.

Gerard, François et al. (2019). *Two-Tier Tax Systems and Firms: Evidence from Brazil.*

Gilovich, Thomas (1983). "Biased evaluation and persistence in gambling." In: *Journal of Personality and Social Psychology* 44.6, pp. 1110–1126. DOI: 10.1037/0022-3514.44.6.1110.

Gordon, Roger and Wei Li (Aug. 2009). "Tax structures in developing countries: Many puzzles and a possible explanation". In: *Journal of Public Economics* 93.7-8, pp. 855–866. DOI: 10.1016/j.jpubeco.2009.04.001.

Haerpfer, Christian et al. (2020). *World Values Survey Wave 7 (2017-2020) Cross-National Data-Set*. en. DOI: 10.14281/18241.1.

Heckman, J. J. and E. J. Vytlacil (Apr. 1999). "Local instrumental variables and latent variable models for identifying and bounding treatment effects". In: *Proceedings of the National Academy of Sciences* 96.8, pp. 4730–4734. DOI: 10.1073/pnas.96.8.4730.

Heckman, James J. and Edward Vytlacil (May 2005). "Structural Equations, Treatment Effects, and Econometric Policy Evaluation1". In: *Econometrica* 73.3, pp. 669–738. DOI: 10.1111/j.1468-0262.2005.00594.x.

Heckman, James J. and Edward J. Vytlacil (2007). "Chapter 70 Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation". In: *Handbook of Econometrics*. Elsevier, pp. 4779–4874. DOI: 10.1016/s1573-4412(07)06070-9.

Imbens, Guido W. and Joshua D. Angrist (Mar. 1994). "Identification and Estimation of Local Average Treatment Effects". In: *Econometrica* 62.2, p. 467. DOI: 10.2307/2951620.

IMF (2011). *Revenue Mobilization in Developing Countries*. Research rep. Fiscal Affairs Department, International Monetary Fund (IMF).

IRS (1996). *Federal Tax Compliance Research, Individual Income Tax Gap Estimates for 1985, 1988, and 1992*. Tech. rep. Publication 1415 (Rev. 4-96)Catalog Number 10263H. Department of the Treasury, Internal Revenue Service.

Keen, Michael (2008). "VAT, Tariffs, and Withholding: Border Taxes and Informality in Developing Countries". In: *Journal of Public Economics*, pp. 1892–1906.

Keen, Michael and Joel Slemrod (Aug. 2017). "Optimal Tax Administration". In: *Journal of Public Economics* 152, pp. 133–142.

Keen, Michael and Stephen Smith (Dec. 2006). "VAT Fraud and Evasion: What Do We Know and What Can Be Done?" In: *National Tax Journal* 59.4, pp. 861–887.

Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken (Feb. 2016). "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors". In: *Quarterly Journal of Economics* 131.1, pp. 219–271.

Kirchler, Erich (2007). *The economic psychology of tax behaviour*. New York: Cambridge University Press. ISBN: 9780511628238.

Kleven, Henrik and Mazhar Waseem (2013). "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Evidence from Pakistan". In: *Quarterly Journal of Economics*, pp. 669–723.

Kleven, Henrik J. et al. (2011). "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark". In: *Econometrica* 79.3, pp. 651–692. DOI: 10.3982/ecta9113.

Kleven, Henrik Jacobsen (Oct. 2016). "Bunching". In: *Annual Review of Economics* 8.1, pp. 435–464. DOI: 10.1146/annurev-economics-080315-015234.

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez (Mar. 2016). "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries". In: *Economica* 83.330, pp. 219–246. DOI: 10.1111/ecca.12182.

Kopczuk, Wojciech and Joel Slemrod (Apr. 2006). "Putting Firms into Optimal Tax Theory". In: *American Economic Review* 96.2, pp. 130–134. DOI: 10.1257/000282806777212585.

Kotchen, M. J. and L. E. Grant (2011). "Does Daylight Saving Time Save Energy? Evidence From a Natural Experiment". In: *The Review of Economics and Statistics*, pp. 1172–1185.

Kowalski, Amanda (June 2016). *Doing More When You're Running LATE: Applying Marginal Treatment Effect Methods to Examine Treatment Effect Heterogeneity in Experiments*. Tech. rep. DOI: 10.3386/w22363.

Liu, Li et al. (Mar. 2021). "VAT Notches, Voluntary Registration, and Bunching: Theory and U.K. Evidence". In: *The Review of Economics and Statistics* 103.1, pp. 151–164. DOI: 10.1162/rest_a_00884.

Maciejovsky, Boris, Erich Kirchler, and Herbert Schwarzenberger (Dec. 2007). "Misperception of chance and loss repair: On the dynamics of tax compliance". In: *Journal of Economic Psychology* 28.6, pp. 678–691. DOI: 10.1016/j.joep.2007.02.002.

Mittone, Luigi (Oct. 2006). "Dynamic behaviour in tax evasion: An experimental approach". In: *The Journal of Socio-Economics* 35.5, pp. 813–835. DOI: 10.1016/j.socec.2005.11.065.

Muralidharan, Karthik and Paul Niehaus (Nov. 2017). "Experimentation at Scale". In: *Journal of Economic Perspectives* 31.4, pp. 103–124. DOI: 10.1257/jep.31.4.103.

Naritomi, Joana (Sept. 2019). "Consumers as Tax Auditors". In: *American Economic Review* 109.9, pp. 3031–3072.

Paula, Aureo de and Jose A. Scheinkman (Oct. 2010). "Value-Added Taxes, Chain Effects, and Informality". In: *American Economic Journal: Macroeconomics* 2.4, pp. 195–221.

Piggott, John and John Whalley (2001). “VAT Base Broadening, Self Supply, and the Informal Sector”. In: *American Economic Review* 91, pp. 1084–94.

Pomeranz, Dina (2015). “No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax”. In: *American Economic Review* 105.8, pp. 2539–2569.

Sabaini, Juan Carlos Gómez and Juan Pablo Jiménez (2012). “Tax structure and tax evasion in Latin America”. In: *Macroeconomics of Development Series* 118.

Saez, Emmanuel (Aug. 2010). “Do Taxpayers Bunch at Kink Points?” In: *American Economic Journal: Economic Policy* 2.3, pp. 180–212. DOI: `10.1257/pol.2.3.180`. URL: `https://www.aeaweb.org/articles?id=10.1257/pol.2.3.180`.

Sarin, Natasha and Lawrence Summers (Nov. 2019). *Shrinking the Tax Gap: Approaches and Revenue Potential*. Tech. rep. DOI: `10.3386/w26475`.

Slemrod, Joel (Jan. 2016). “Tax Compliance and Enforcement: New Research and its Policy Implications”. In: *Ross School of Business Working Paper*. University of Michigan. URL: `http://ssrn.com/abstract=2726077`.

— (2019). “Tax Compliance and Enforcement”. In: *Journal of Economic Literature* 57.4, pp. 904–954.

Slemrod, Joel and C Gillitzer (2014). *Tax Systems*. Cambridge, Massachusetts, USA: The MIT Press.

Slemrod, Joel, Obeid Ur Rehman, and Mazhar Waseem (Apr. 2020). “Pecuniary and Non-Pecuniary Motivations for Tax Compliance: Evidence from Pakistan”. In: *Review of Economics and Statistics*, Accepted.

Sow, Soule and Mesay Gebresilasse (Sept. 2020). “Effect of VAT Adoption on Manufacturing Firms in Ethiopia”. In: *International Journal of Economics and Finance* 12.10, p. 75. DOI: `10.5539/ijef.v12n10p75`.

Taubinsky, Dmitry and Alex Rees-Jones (Aug. 2016). *Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment*. Tech. rep. DOI: `10.3386/w22545`.

Virmani, Arvind (1989). “Indirect Tax Evasion and Production Effeciency”. In: *Journal of Public Economics* 39, pp. 223–237.

Waseem, Mazhar (2018). “Taxes, Informality and Income Shifting: Evidence from a Recent Pakistani Tax Reform”. In: *Journal of Public Economics*, pp. 41–77.

— (Mar. 2019). “Does Cutting the Tax Rate to Zero Induce Behavior Different from Other Tax Cuts? Evidence from Pakistan”. In: *Review of Economics and Statistics*, Accepted.

Waseem, Mazhar (Jan. 2020). "The Role of Withholding in the Self-Enforcement of a Value-Added Tax". In: *Review of Economics and Statistics*, Revise and Resubmit.

— (2020b). *Overclaimed Refunds, Undeclared Sales, and Invoice Mills: Nature and Extent of Noncompliance in a Value-Added Tax*. Tech. rep.

Yithzaki, S. (1974). *A note on Income Tax Evasion: A Theoretical Analysis*. Journal of Public Economics, pp. 201–202.

Vita

Jawad Shah

Education

MPA, Maxwell School, Syracuse University, 2013
B.Sc. in Electrical Engineering, University of Engg. & Technology Lahore, 2002
B.A. Mathematics & History, University of the Punjab, 1999

Employment

Additional Collector, Federal Board of Revenue (FBR) Pakistan, 2002- current
Research Assistant, University of Kentucky, 2016-2021

Awards

Fulbright Master Scholarship for studies at Syracuse University, 2012-2013
Kentucky Opportunity Fellowship, 2015-2016
Quaid-e-Azam scholarship and Silver Medalist, HSSC Examination, 1997